Spatiotemporal Analysis of Magnetic Resonance Perfusion Data: The Multi-Compartment Problem

Eve Sophia Shalom

University of Leeds

School of Physics and Astronomy

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

November, 2023

Intellectual Property Statement

The candidate confirms that the work submitted is their own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Joint Publications

Work from the following jointly authored papers, included as (Shalom *et al.*, 2024a), (Shalom *et al.*, 2024b) and (Shalom *et al.*, 2024c) in the References, are included within this thesis. The candidate will be referred to as E.S. Shalom for clarity.

Journal Papers:

E.S. Shalom, A. Khan, S. Van Loo, and S.P. Sourbron. Current Status in Spatiotemporal Analysis of Contrast-Based Perfusion MRI. Magnetic Resonance in Medicine. 2024; 91: 1136-1148. doi: 10.1002/mrm.29906

Contribution: *All literature review and manuscript drafting completed by E.S. Shalom. Comments and edits were iterated with co-authors.* Chapter 4 includes work contained in this paper.

E.S. Shalom, H. Kim, RA. van der Hijden, Z. Ahmed, R. Patel, D.A. Hormuth II, J.C. DiCarlo, T.E. Yankeelov, N.J. Sisco, R.D. Dortch, A.M. Stokes, M. Inglese, M. Grech-Sollars, N. Toschi, P. Sahoo, A. Singh, S.K. Verma, D.K. Rathore, A.S. Kazerouni, S.C. Partridge, E. LoCastro, R. Paudyal, I.A. Wolansky, A. Shukla-Dave, P. Schouten, O.J. Gurney-Champion, R. Jiřík, O. Macíček, M. Bartoš, J. Vitouš, A.B. Das, S. Gene Kim, L. Bokacheva, A. Mikheev, H. Rusinek, M. Berks, P.L. Hubbard Cristinacce, R.A. Little, S. Cheung, J.P.B. O'Connor, G.J.M. Parker, B. Moloney, P.S. LaViolette, S. Bobholz, S. Duenweg, J. Virostko, H.O. Laue, K. Sung, A. Nabavizadeh, H.S. Rad, L.S. Hu, S.P. Sourbron, L.C. Bell, and A.F. Kazerooni. The ISMRM Open Science Initiative for Perfusion Imaging (OSIPI): Results from the OSIPI–Dynamic Contrast-Enhanced challenge. Magnetic Resonance in Medicine. 2024; 91: 1803-1821. doi: 10.1002/mrm.29909

Contribution: *E.S. Shalom carried out all design and simulation of synthetic challenge data. E.S. Shalom implemented and carried out analysis and presentation of all submitted challenge data in Python. First draft of manuscript by E.S. Shalom (Methods: [Synthetic Data, Leaderboard Evaluation, Statistical Analysis], Results, Discussion: [Accuracy, Repeatability, Implication on Future Challenge Design, Study Limitations]) and A.F. Kazerooni (Introduction, Methods: [Challenge Setup, Clinical Data, Tumour*

Segmentation], Discussion: [Reproducibility, Implication on Future Challenge Design, Study Limitations]). Further draft edits were iterated between E.S. Shalom and A.F. Kazerooni. Other co-authors provided edits and comments. Contribution from other co-authors include submission of data to be analysed (entry teams), replication of submission data (neutral evaluators), challenge testing, collation of clinical data, and challenge conception and advertisement (OSIPI Taskforce 6.2 - Co-led by A.F. Kazerooni and H. Kim). Chapter 3 includes work contained in this paper.

E.S. Shalom, S. Van Loo, A. Khan, and S.P. Sourbron. Identifiability of spatiotemporal tissue perfusion models. Physics in Medicine and Biology. 2024. doi: 10.1088/1361-6560/ad4087.

Contribution: *All simulation codes designed and written by E.S. Shalom. Data analysis and presentation as well as manuscript drafting completed by E.S. Shalom. The co-authors are the supervisory team who provided guidance and ideas for progression, as well as comments and edits to the manuscript.* Chapter 5 includes work contained in this paper.

Conference Abstracts:

All conference abstracts associated are listed here, the candidate was the presenting author in all cases.

E.S. Shalom, S.P. Sourbron, A.F. Kazerooni, S. Van Loo, and A. Khan. (2021) How to fake a brain scan: A digital reference object for testing DCE-MRI analysis methods. Physics Postgraduate Research Symposium 2021, University of Leeds. e-Poster Presentation.

E.S. Shalom, S. Van Loo, A. Khan, and S.P. Sourbron. Tracer kinetic field theory: Towards spatiotemporal recovery in two compartment systems. (2022) ISMRM Workshop on Perfusion MRI: From Head to Toe. Poster presentation.

E.S. Shalom, S. Van Loo, A. Khan, and S.P. Sourbron. (2022) A Proof of Concept Study: Towards recovery of spatial kinetic parameters in two compartment systems. Postgraduate Symposium of the British Chapter of ISMRM 2022, London, UK. Poster presentation.

E.S. Shalom, H. Kim, RA. van der Hijden, Z. Ahmed, R. Patel, D.A. Hormuth II, J.C. DiCarlo, N.J. Sisco, R.D. Dortch, A.M. Stokes, M. Inglese, M. Grech-Sollars, N. Toschi, P. Sahoo, A. Singh, S.K. Verma, D.K. Rathore, A. Nabavizadeh, H.S. Rad, L.S. Hu, S.P. Sourbron, L.C. Bell, and A.F. Kazerooni. (2022) The Open Science Initiative for Perfusion Imaging (OSIPI): Results from the DCE-MRI challenge. ISMRM 2022, London, UK. Oral presentation.

E.S. Shalom, S.P. Sourbron, A. Khan, and S. Van Loo. (2023) Recovering Kinetic Parameters in Spatial Two Compartment Systems: Proof of Concept. Postgraduate Symposium of the British Chapter of ISMRM 2023, Manchester, UK. Poster presentation.

Acknowledgements

Firstly, I would like to thank my supervisors Steven Sourbron, Sven Van Loo and Amirul Khan for all their advice and invaluable help from fruitful discussions and helpful feedback to frustrating hours spent debugging code. Particularly, I am grateful to Steven for the generous amounts of time and knowledge of the field he has provided over the years.

I would like to also acknowledge EPSRC for providing the funding for this CASE studentship alongside Bayer AG the industry partner.

Thank you also to Anahita Fathi-Kazerooni, Harrison Kim, Laura Bell and the rest of OSIPI Taskforce 6.2 for the great collaboration that went into the DCE-Challenge and our resulting paper[1], it was truly a pleasure to get to work with you all.

To Julia, Jacob, Greg, Rob and Hannah, I'd like to thank you for being the most amazing group of housemates I could have asked for. I thank you all for providing sage advice, fun days and nights out, alongside much hilarity and support when it was needed most.

To the whole Astronomy office past and present, thank you all for the numerous tea breaks, lunch times and Friday evenings spent fondly at the Fenton. We have truly been on a journey over the past four years, physically from our old spot in EC Stoner to the penthouse of the Bragg Building - but also through the depths of the pandemic with the virtual pub a welcome catch-up when we couldn't meet physically. Even though I am not an Astronomer, everyone always made me feel at home - and I have even picked up a startling amount of jargon!

Thank you also to my long-standing physics family from my Birmingham undergraduate days, I always looked forward to the fun weekends away catching up with you all regardless of how work was going.

To my partner Ollie, thank you for your continued love and support. I am thankful for all the inspiring bird memes and dinosaur

[1] See contribution details in "Joint Publications" on ii.

facts you have sent to me and am forever amused by your intense and ongoing enthusiasm for the passingly mentioned Cuttlefish algorithm.

Lastly to my family, thank you to Mum, Dad, and Luke for your everlasting love and support. Thanks for letting me ramble about tracer kinetics and optimisation when needed. I am forever grateful to you three for keeping me enthusiastic and motivated towards the end.

For Mum, Dad and Luke.

“ The beginning is always today. ” – Mary Wollstonecraft Shelley

Abstract

Quantitative analysis in perfusion magnetic resonance imaging (MRI) has shown real potential for extracting clinically relevant diagnostic information. Integration of quantitative perfusion MRI into clinical care is limited, in part due to issues in parameter reproducibility between centres. An assessment of 10 typical dynamic contrast-enhanced MRI analysis software packages is presented across clinical and synthetic datasets in terms of accuracy, repeatability, and reproducibility. Results show differences in K^{trans} quantification between the 10 software packages dependent on multiple methodology choices. While reduction of manual methodology steps and improved documentation is recommended, development of spatiotemporal modelling is proposed as a potential paradigm change to improve perfusion analysis. A literature review of early contributions within spatiotemporal modelling for perfusion MRI is carried out, identifying the multi-compartment problem as the most critical and open challenge.

A novel gradient descent-based algorithm is proposed which uses both one- and two-compartment toy-models to determine system identifiability. The whole 1D spatial system is fitted at once from *in-silico* data. Results across 3 systems for each model showed identifiability of transport and system influx for one-compartment systems. Multiple solutions were reported in two-compartment systems. The long computational times observed restricted the extension of this approach to higher dimensional systems.

The suitability of Physical Informed Neural Networks (PINNs) for spatiotemporal modelling in perfusion MRI has been investigated as a solution for decreased computational times. Implementation of PINNs reports similar outcomes in one- and two-compartment systems compared with the standard approach. Boundary condition enforcement in two-compartment systems is recommended in future developments. Results showed decreased computational time compared to standard optimisation, with straightforward extension into higher spatial dimensions.

The proposed algorithms show spatiotemporal one-compartment models are readily identifiable from 1D toy-models. This work reports multiple solutions for two-compartment systems, indicating further work on fundamental identifiability and optimisation strategies is required.

Contents

List of Figures

List of Tables

Acronym Glossary

ADC	Apparent Diffusion Coefficient
AIF	Arterial Input Function
ASL	Arterial Spin Labelling
AUC	Area Under Curve
BAT	Bolus Arrival Time
CA	Contrast Agent
CFD	Computational Fluid Dynamics
CNNs	Convolutional Neural Networks
CNR	Contrast to Noise Ratio
DCE	Dynamic Contrast Enhanced
DSC	Dynamic Susceptibility Contrast
FID	Free Induction Decay
MRI	Magnetic Resonance Imaging
MTT	Mean Transit Time
NMR	Nuclear Magnetic Resonance
NNs	Neural Networks
NSF	Nephrogenic System Fibrosis
1C	One Compartment Model
PINNs	Physics Informed Neural Networks
PK	Pharmacokinetic
RF	Radio Frequency
ROI	Region of Interest
SNR	Signal to Noise Ratio
SOP	Standard Operating Procedure
SPGR	Spoiled Gradient Echo
TCIA	The Cancer Imaging Archive
3C	Three Compartment Model
TK	Tracer Kinetic
tkPINN	Tracer Kinetic Physics Informed Neural Network
TM-ROI	Tumor mask ROI
TTP	Time to Peak
2C	Two Compartment Model
2CXM	Two Compartment Exchange Model
VFA	Variable Flip Angle

Parameter Glossary

$\mathbf{m}$	Magnetic Moment
$\mathbf{S}$	Spin Angular Momentum
$\mathbf{B}$	Applied External Magnetic Field
B_0	Applied External Magnetic Field Magnitude Along z-Direction
γ	Gyromagnetic Ratio
ω	Larmor Frequency
$\mathbf{M}$	Macroscopic Magnetisation
M_{xy}	Macroscopic Transverse Magnetisation
M_z	Macroscopic Longitudinal Magnetisation
M_0	Steady State Macroscopic Magnetisation
T_1	Longitudinal Relaxation Time
T_{10}	Pre-contrast Longitudinal Relaxation Time
T_2	Transverse Relaxation Time
$T_2^\star$	Apparent Measured Transverse Relaxation Time
T_{20}	Pre-contrast Transverse Relaxation Time
$T_{20}^\star$	Apparent Measured Pre-contrast Transverse Relaxation Time
R_1	Longitudinal Relaxation Rate
R_{10}	Pre-contrast Longitudinal Relaxation Rate
R_2	Transverse Relaxation Rate
$R_2^\star$	Apparent Measured Transverse Relaxation Rate
R_{20}	Pre-contrast Transverse Relaxation Rate
$R_{20}^\star$	Apparent Measured Pre-contrast Transverse Relaxation Rate
r_1	Longitudinal Relaxivity
r_2	Transverse Relaxivity
T_R	Repetition Time
T_E	Echo Time
A_0	Scanner Dependent Scaling Term
ϕ	Flip Angle
S	Signal
S_0	Pre-contrast Signal
c^q	Local Concentration in Compartment q
C^q	Tissue Concentration for Compartment q
C	Total Tissue Concentration
C^{meas}	Measured Total Tissue Concentration
C^{pred}	Predicted Total Tissue Concentration
C^{max}	Maximum Total Tissue Concentration
C^{Net}	Network Predicted Total Tissue Concentration

ΔC_{rel}	Percentage Error Relative to Maximum Tissue Concentration
c_a or c_{in}	Arterial Input Function
c_l	2D System Left Boundary Local Concentration
c_r	2D System Right Boundary Local Concentration
c_t	2D System Top Boundary Local Concentration
c_b	2D System Bottom Boundary Local Concentration
F_p	Plasma Flow
PS	Permeability-Surface Area Product
K^{trans}	Volume Transfer Constant
k_{ep}	Reflux Rate
v^q	Volume Fraction of Compartment q
v	Total Volume Fraction
λ^a	Ratio of Arterial Volume Fraction to Total Volume Fraction
$\vec{f}^q$	Surface Flow Density in Compartment q
$\vec{j}$	Surface Flux
$\vec{u}^q$	Velocity in Compartment q
$D^{q,\star}$	Molecular Diffusion Coefficient in Compartment q
D^q	Pseudo Diffusion Coefficient in Compartment q
F	Perfusion Exchange (Arterial to Venous)
K^{va}	Arterial to Venous Exchange Rate
$k_{i,i-1}$	Exchange Rate Constant From Voxel $i-1$ into i
J	Concentration Influx
^{a}	Arterial Compartment
^{v}	Venous Compartment
^{e}	Interstitial Compartment
^{p}	Plasma Compartment
x	x-Direction Spatial Coordinate
Δx	Voxel Width
y	y-Direction Spatial Coordinate
Δy	Voxel Height
t	Temporal Coordinate
Δt	Simulation Timestep
N	Total Number of Voxels
N_x	Total Number of Voxels in x-Direction
N_y	Total Number of Voxels in y-Direction
T	Total Number of Time Points
P_{rec}	Recovered Parameters
P_{gt}	Ground Truth Parameters
E_{rel}	Percentage Error Relative to Parameter Mean Value
$\mathfrak{X}$	Co-variance Matrix
$\mathbf{H}$	Hessian Matrix
$\mathfrak{J}$	Jacobian Matrix
ρ	Correlation Coefficient Matrix
$\mathcal{L}$	Total Loss Function
$\mathcal{L}_{Physics}$	Physics Loss Function
$\mathcal{L}_{Data}$	Data Loss Function
Φ	Residual Equation

Chapter 1

Introduction

1.1 Context and Motivations

This thesis investigates different aspects relevant to perfusion Magnetic Resonance Imaging (MRI). Perfusion MRI techniques include Dynamic Contrast-Enhanced MRI (DCE-MRI), Dynamic Susceptibility-Enhanced MRI (DSC-MRI), and Arterial Spin Labelling (ASL). For DCE and DSC-MRI intravascular indicators are introduced into patients, while magnetically labelled water is applied in ASL. Applications of perfusion MRI are aimed towards clinical practice and drug development.

Over the years, analysis of both contrast-based (DCE and DSC-MRI) and ASL techniques have developed from inherently subjective qualitative methods to more objective quantitative measures. One of the core aims for analysis in any branch of medical imaging is actionable patient outcomes, via the use of biomarkers which can be reliably applied across different centres (O'Connor *et al.*, 2017).

Biomarkers are defined characteristics which can be measured quantitatively as an assessment of usual biological or pathogenic processes (FDA-NIH Biomarker Working Group, 2016). The interest in developing biomarkers for clinical use stems from the objectivity of quantitative measures. Examples of patient outcomes from non-imaging biomakers are numerous, with urine and blood tests routinely used to assess sugar, protein or hormone levels to aid the diagnosis of conditions such as diabetes and kidney disease. Equal value could be contributed to clinical decisions from the development of imaging-based biomarkers which have the ability to deduce values of interest *in-vivo*.

A potential biomarker that has been of interest within the DCE community is the transfer rate of contrast agent (CA) from the blood plasma into the extracellular extravascular space, usually referred to as the volume transfer constant or K^{trans} parameter (Dickie *et al.*, 2023). K^{trans} has been shown to be a useful quantitative parameter in several applications, such as tumour grading (Zhang *et al.*, 2012) and assessment of tumour treatment response (Kickingereder *et al.*, 2015). Despite several decades of development, advanced quantitative analysis of perfusion MRI is very limited in standard clinical practice. This is due in part to issues in standardisation between centres and research teams (Kim, 2018). Even when based on the same underlying models and physiological assumptions, different software implementations or algorithms can return different values for target parameters such as K^{trans} (Beuzit *et al.*, 2016).

Development of analysis techniques by incorporating spatiotemporal approaches to increase modelling sophistication could allow for more accurate extraction of parameters such as perfusion and blood flow velocity. Access to such parameters and any heterogeneity across tissues would be invaluable for future biomarker developments in order to report meaningful results where dysfunction or disease are not homogeneous. For example, a spatial assessment of liver function could help offer liver surgery to patients with underlying liver disease who are currently ruled out by default (Elsharif *et al.*, 2021). Knowledge of the heterogeneity behind overall liver functionality would help inform risk assessments and potentially allow more patients access to this curative option (Elsharif *et al.*, 2021).

This thesis will focus on the motivation for moving towards spatiotemporal models, away from current standard methods in perfusion MRI. Additionally, this thesis aims to contribute to the development of inverse approaches for spatiotemporal models and investigate their application on synthetic test systems.

1.2 Overview of Thesis Aims

The overall focus of this thesis is to provide insight into several areas that could contribute to the progression of quantitative analysis in perfusion MRI. This work has four main aims (Aims I-IV), as outlined below.

Aim I: Evaluate K^{trans} quantification in current DCE-MRI perfusion analysis software.

Quantitative analysis techniques for DCE-MRI have been incrementally developed over the past 30 years (Lohrke *et al.*, 2016). However, it is recognised that varied implementation approaches can influence modelling outcomes (O'Connor *et al.*, 2007; Beuzit *et al.*, 2016). To this end the thesis outlines an investigation of K^{trans} quantification from 10 current software packages in the context of brain tumours. This was an international collaboration project,[1] where 10 entry teams submitted K^{trans} values from analysis of one standard set of data. The 10 packages are assessed in terms of accuracy, repeatability and reproducibility. From this evaluation it is shown there are large differences in reported values between software packages. This variation suggests a step change is needed, either from further standardisation work to ensure cross-centre parameter agreement or development of more advanced modelling approaches.

Aim II: Review previous contributions of spatiotemporal modelling in perfusion MRI.

Current standard analysis considers voxels in isolation and utilise a global input. In contrast, advanced modelling accounting for the spatiotemporal nature of perfusion treats scanning voxels as a set of interacting spaces. This, therefore, is a much more realistic framework for transport processes, such as convection, diffusion, perfusion and cellular uptake within the body (Sourbron, 2014). Historically, spatiotemporal modelling has been inaccessible due to the complex nature of the problem. This thesis explores the current state of the art work within spatiotemporal modelling for perfusion MRI. While work within

[1] Please see "Joint Publications" on Page ii for contribution details.

the field has previously been sparse (Thacker *et al.* (2003); Frank *et al.* (2008); Pellerin *et al.* (2007)) there is increased interest in development of novel techniques (Sinno *et al.*, 2021, 2022; Zhou *et al.*, 2021; Zhang *et al.*, 2022, 2023; Liu *et al.*, 2021a; Naevdal *et al.*, 2016), after the introduction of a comprehensive framework (Sourbron, 2014). The literature review carried out within this thesis identifies that recovery of one-compartment systems has been the main focus of previous studies, with extremely limited implementation of multi-compartmental systems (Naevdal *et al.*, 2016).

Aim III: Investigate identifiability of one- and two-compartment spatiotemporal models.

In order to address the current lack of multi-compartment implementations, this thesis develops and tests a gradient descent-based optimisation method for simple one- and two-compartment spatiotemporal toy-models. These investigations show there may be system identifiability issues with multiple solutions found to provide acceptable data fitting. Long computational times were also needed, even for these small systems. To make analysis of clinically relevant datasets feasible, this work identified a need for more efficient optimisation approaches.

Aim IV: Assess Physics Informed Neural Networks for suitability in scaling up spatiotemporal methods.

Physics Informed Neural Networks (PINNs) (Raissi *et al.*, 2019) have been used in both forwards modelling and inverse approaches, by constraining outputs to adhere to defined physical laws. These are attractive due to their generalised setup and ability to handle large volumes of data efficiently. This thesis explores the use of PINNs within spatiotemporal perfusion modelling as a potential solution to reduce computational time when compared to standard optimisation methods. An initial PINN is designed and applied to the simple one- and two-compartment spatiotemporal toy-models previously investigated. The investigations showed promising outcomes and raise important questions to be addressed in future work.

1.3 Thesis Organisation

This thesis presents a study into the development of perfusion MRI analysis as detailed above in Section 1.2. The organisation of chapters within this thesis is detailed below.

Chapter 2 comprises a background of relevant topics to equip the reader with vital information key to the rest of the thesis matter. This consists of a brief overview of MRI, DCE-MRI, standard analysis techniques of DCE-MRI data, spatiotemporal theory for DCE-MRI analysis and PINNs.

Chapter 3 presents the reported K^{trans} values across 10 software packages applying typical single-voxel tracer kinetics (TKs). These software packages are applied to the same clinical and synthetic data sets with metrics calculated to assess the accuracy, repeatability and reproducibility of each package.

Chapter 4 details an in-depth review, lacking in the current literature, of the current advances in the spatiotemporal TK field. This chapter gives a coherent description of models and implementation details in terms of the transport mechanisms at work and the number of compartments addressed.

Chapter 5 details a proof of concept study applying a gradient descent-based method for the inversion problem for both one- and two-compartment spatiotemporal systems. This chapter investigates the feasibility of recovering spatiotemporal parameters from DCE-MRI data using in-silico toy-models. This is particularly centred on the reconstruction of two-compartment systems without the need for additional constraints or assumptions which is currently absent from the literature landscape. The resulting parameter retrievals are illustrated for noiseless, noisy and under-sampled systems.

Chapter 6 presents a novel application of a PINNs-based approach for multi-compartment spatiotemporal systems modelled via TKs. This chapter builds on the work in Chapter 5 by utilising the same toy-models. The resulting parameter recovery is illustrated for noiseless, noisy and under-sampled systems, showing a varied response. An initial investigation into a 2D one compartment system is also presented.

Chapter 7 provides a summary of this thesis with conclusions of the work

presented, and details recommendations for future work within the perfusion imaging community.

Chapter 2

Background

This Chapter summarises the relevant background material for this thesis. This includes an overview of DCE-MRI, single voxel analysis techniques of DCE-MRI data, spatiotemporal theory for DCE-MRI analysis and physics-informed neural networks.

2.1 Magnetic Resonance

Both quantum and classical mechanics may be used to build a descriptive picture of the Nuclear Magnetic Resonance (NMR) phenomenon. Adequate description and interpretation of the phenomena in MRI is possible using the classical mechanics treatment, and will be considered here. An intuitive explanation of the theory with consistency between quantum and classical mechanics formalism can be found in Levitt (2013, Chp. 2 and 10) and Hanson (2015), with Hanson (2008) addressing common myths that persist in the topic. The overview of MRI principles included in this section summarises ideas from Levitt (2013, Chp. 2), Hanson (2015), and Pickens (2000).

2.1.1 Individual Magnetic Moments

When presenting the underlying physics of NMR it can be useful to appreciate the quantum description of atomic nuclei. This was detailed by Dirac in the 1930s, which predicted an intrinsic angular momentum known as spin.

A nucleus also has an intrinsic magnetic moment[1], $\mathbf{m}$, related proportionally

[1] A magnetic moment, $\mathbf{m}$, is the strength and orientation of the magnetic field associated with an object. The magnetic moment is employed to relate the torque, τ, experienced by an object in an external field, B_0 by the cross product $\tau = \mathbf{m} \times B_0$.

to the spin angular momentum, $\mathbf{S}$, by the gyromagnetic ratio, γ:

$$\mathbf{m} = \gamma \mathbf{S}. \tag{2.1}$$

The direction of the magnetic moment is either positioned along, $\gamma > 0$, or in the opposite direction, $\gamma < 0$, to the spin. The direction of the spin angular momentum will be referred to by the spin polarisation axis following from the description in Levitt (2013).

In a system unaffected by external magnetic fields the magnetic moments are distributed isotropically, with all orientations equally likely (Levitt, 2013). A representation of isotropic spin polarisation axes is shown in Figure 2.1. When an external magnetic field, B_0, is applied the nuclei spin polarisation axis precesses around the field direction. The relation stated in (2.2) describes a nuclei with a gyromagnetic ratio, γ, precessing in the applied field, B_0, at the Larmor frequency, ω (Levitt, 2013).

$$\omega = -\gamma B_0 \tag{2.2}$$

Here γ is a characteristic of the nucleus dependent on mass, charge and spin, allowing calculation of the Larmor frequency which describes the number of nucleus precessions per second within a magnetic field. This precession is a consequence of the angular momentum property of spin coupled with the torque experienced by the magnetic moment. The idea is analogous to a spinning top precessing about the gravitational field direction. Precession describes a cone like shape, only varied by the initial angle of the spin polarisation to the applied field direction, shown in the middle pane of Figure 2.1. Very weak local variation in magnetic field causes a slight wandering effect in the precession angle. This effect is not isotropic with a higher probability to drift towards a lower magnetic energy state where the spin polarisation axis and field direction are aligned (Levitt, 2013).

It is essential to understand that this anisotropic distribution is in thermal equilibrium, with individual spins precessing and wandering continuously. This distribution gives rise to the macroscopic net magnetisation. If a sample has a magnetic field instantaneously applied, the net magnetisation grows along the

field direction from zero through this "biased wandering motion" (Levitt, 2013).

2.1.2 Macroscopic Precession of Bulk Magnetisation

For a classical description, nuclei are described by small bar magnets. When exposed to an external magnetic field there is a tendency to align along this field. For a system at a finite temperature thermal agitation constantly changes the alignment of these bar magnets, but a preference along the external field remains (Hanson, 2008). This produces a net magnetisation in a sample of nuclei as shown in Figure 2.1.

Each nucleus behaves with quantum nature and exist in a superposition of spin states while unobserved. Conducting a measurement on a single spin-$\frac{1}{2}$ nucleus results in collapse to one of the two eigenstates; either parallel or anti-parallel to the applied magnetic field. These spin states correspond to two energy states, where 'spin-up' (parallel) is more energetically favourable. During MRI many nuclei are measured simultaneously for their bulk properties so single nuclei are not forced into this measurement state, as laid out in the Appendix of Hanson (2008). This gives rise to the representation shown in Figure 2.1 which is a helpful visualisation of both the quantum and classical description (Levitt, 2013; Hanson, 2008).

Response of protons in the body to external magnetic fields can be fully described by the macroscopic Bloch Equations proposed in 1946 (Bloch *et al.*, 1946). These phenomenological relationships describe nuclear magnetism and the response to externally applied fields. The full Bloch equation is defined by (2.3) in terms of the magnetisation, $\mathbf{M}$, external magnetic field, $\mathbf{B}$, relaxation times T_1 and T_2.

$$\frac{d\mathbf{M}(t)}{dt} = \gamma(\mathbf{M}(t) \times \mathbf{B}(t)) - \frac{\mathbf{M_{xy}}(t)}{T_2} + \frac{M_0 - M_z(t)}{T_1}\hat{\mathbf{k}} \tag{2.3}$$

Where γ is the gyromagnetic ratio of the nuclei, M_0 describes the magnetisation in the steady state ($t \rightarrow \infty$), and M_z and $\mathbf{M_{xy}}$ are the longitudinal and transverse components of the magnetisation, respectively. These relations form the basis of MRI. For the special case of free relaxation,

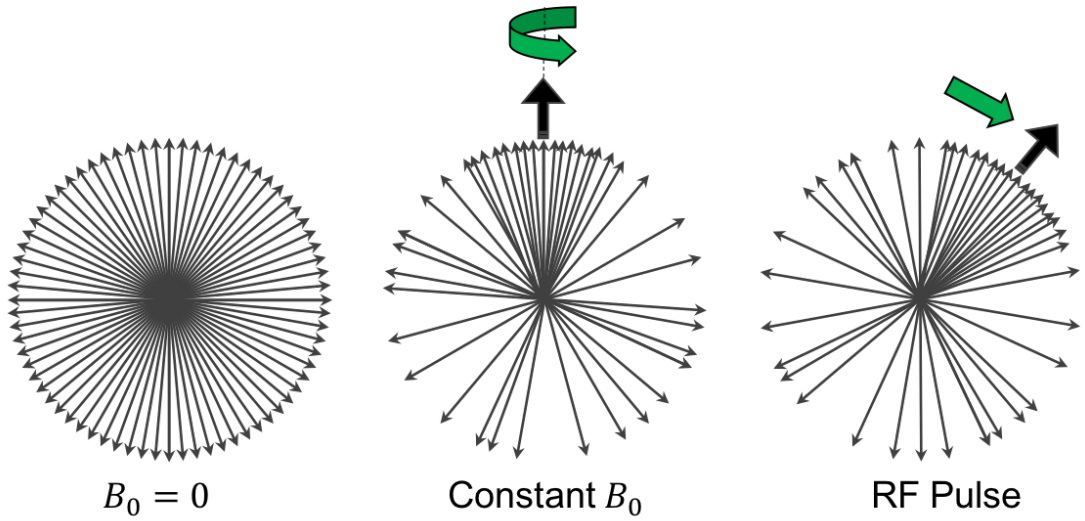

Figure 2.1: Useful representation of the spin distribution in various external conditions. *Left:* no external magnetic field. *Middle:* static external magnetic field in the upwards direction, with precession shown by the green arrow. The bulk magnetisation shown by the wide vertical black arrow. *Right:* The rotation of the spin distribution from an RF pulse on resonance.

(2.3) may be simplified with the relaxation of the magnetisation components in a static magnetic field described by (2.4) and (2.5) (Pickens, 2000).

$$M_z(t) = M_z(0)e^{-\frac{t}{T_1}} + M_0(1 - e^{-\frac{t}{T_1}}) \tag{2.4}$$

$$M_{xy}(t) = M_{xy}(0)e^{-\frac{t}{T_2}} \tag{2.5}$$

2.2 Magnetic Resonance Imaging

The independent works of Bloch *et al.* (1946) and Purcell *et al.* (1946) detailed the discovery of magnetic resonance processes in bulk matter (Pickens, 2000, p.375). These and future works on NMR led to the emergence of MRI, a medical imaging technique exploiting the same nuclei behaviour. Relaxation time differences between healthy tissue and tumours shown by Damadian (1971) highlighted promising medical applications. The use of NMR for imaging purposes was first shown by Lauterbur (1973) where water tubes were imaged. Notably, Mansfield & Maudsley (1977) presented the first cross sectional image of a human hand, which they remarked revealed "considerable anatomical detail, particularly of the soft tissue regions".

Since these early developments, including the first body scanner (Damadian *et al.*, 1977), many types of MRI have evolved to exploit different aspects of nuclei behaviour during NMR. Examples include DCE-MRI which utilises the effect of CAs on signals to extract perfusion information, and Diffusion Weighted Imaging which uses differences in water diffusion to create images.

2.2.1 Excitation Pulse

Relaxation processes, defined by (2.4-2.5), may be investigated by system excitation with a radio frequency (RF) pulse. When a resonant circularly polarised RF pulse is applied, the net magnetisation tilts away from the equilibrium position along B_0. With the net magnetisation no longer aligned along B_0, a torque force produces a precession of the net magnetisation at the Larmor frequency (Hanson, 2008).

As soon as the RF pulse is switched off, the protons in the body will undergo free relaxation processes detailed in (2.4-2.5). The longitudinal component, M_z, exponentially grows back to its maximal equilibrium value (2.4) characterised by the T_1 relaxation time. This T_1 relaxation captures the effect of the "biased wandering motion" described in Section 2.1.1. Simultaneously, the transverse component, M_{xy}, undergoes an exponential decay (2.5) characterised by the T_2 relaxation time.

The reduction in the transverse component is due to the dephasing of spins relative to each other and direct interactions between protons. This is a consequence of variation in the immediate local environment of the protons, resulting in a distribution of precessional frequencies. This variation is a combination of purely local effects from the unique physical environment (T_2) and also imperfection of the applied field (T_2'), each causing a relaxation effect (Pickens, 2000). As both T_2 and T_2' relaxation affect the transverse component, the combined effect $T_2^\star$ is measured. The free induction decay (FID) detected by the receiver is created by the precession of the net magnetisation in the transverse plane (van Geuns *et al.*, 1999). The relaxation times T_1, $T_2^\star$ and T_2 may be extracted from FID results.

2.2.2 Spatial Encoding

Gradient fields are used to encode spatial information into signals produced from T_1 and T_2 relaxation processes. A field gradient along the body varies the Larmor frequency, allowing selective excitation of a slice using a narrow RF pulse (Pickens, 2000). Frequency and phase encoding gradients are used to localise signals within the slice. The phase encoding gradient is applied between excitation and detection for a short time period, producing additional phase differences proportional to position. Repetition of the process with a varying time period collects signals with different phase encoding (van Geuns *et al.*, 1999). Frequency encoding is used to differentiate between identical phase encoding, with the gradient applied during the detection of signals. Combination of phase and frequency information allow a grid of unique points to be produced, called the K-space. A 2D Fourier transform is applied to process this raw data into spatial intensities. The details of spatial encoding discussed here have been summarised from van Geuns *et al.* (1999).

2.3 Dynamic Contrast Enhanced MRI

The notion that paramagnetic ions could modify proton relaxation times was proposed in the late 1970s (Lauterbur *et al.*, 1978), just after the development of the first body MRI scanners (Damadian *et al.*, 1977). These concepts have grown into the technique of DCE-MRI, which is widely applied for clinical purposes (Lohrke *et al.*, 2016). The field of DCE-MRI has grown hugely since the first approved Gadolinium-based CAs became available. A detailed overview of key developments in DCE-MRI can be found in Lohrke *et al.* (2016), with advances in organ specific techniques highlighted. The change in T_1 over a tissue allows a dynamic output, as the paramagnetic ions are distributed throughout the body through perfusion processes. A review of perfusion imaging principles and techniques may be found in Jahng *et al.* (2014).

A broad group of CAs are employed across imaging modalities to provide enhanced image information. These CAs fall into two distinct groups: tracers

and indicators. Historic definitions of "foreign indicators" and "native tracers" in Perl *et al.* (1975) define these as introduced substances which are detectable: "foreign indicators" are defined as chemically different to any systemic substance, with "native tracers" being chemically equivalent to a systemic substance. For example, magnetically labelled water in ASL would be considered a tracer, and Gadolinium-based CAs an indicator. Generally, these terms are used interchangeably within the field as the modelling is equivalent when the systemic substances and indicators undergo the same transport processes (Bergner *et al.*, 1967). This thesis follows the general convention within the field which names these models as TK modelling regardless of CA type.

2.3.1 Safety Concerns

Generally, Gadolinium-based CAs are considered safe but some safety concerns are noteworthy. The major concerns are of Gadolinium retention and nephrogenic systemic fibrosis (NSF).

Gadolinium retention is a phenomenon where the molecules from the CA remain within the patient. Although Gadolinium-based enhancement can provide valuable information, the clinical impact of Gadolinium retention is unclear and remains under monitoring (Al-Muhanna, 2022).

NSF is usually observed in patients with already compromised kidney function and is a disease characterised by the hardening of connective tissues (Schieda *et al.*, 2019). Inability to evacuate Gadolinium agents effectively can act as a trigger for NSF (Shah & Olivero, 2017). As such, for these compromised patients indicator-based investigations are not usually recommended. However, studies using newer iterations of Gadolinium agents show this recommendation could be revised after further investigation (Bhargava *et al.*, 2021).

2.3.2 Conversion from Signal to Concentration

Direct measurement of CA concentration is not possible in MRI and must be calculated from the resulting effect on relaxation times. As discussed in

Section 2.3 the introduction of a CA has a resultant impact on the signal acquired within a scan. In order to utilise these signal changes the concentration of the CA must be extracted.

In the low-concentration regime, it is usually valid to assume a linear relationship between concentration and relaxation time (Pintaske *et al.*, 2006). Pre-contrast relaxation times $(T_{10}, T_{20}^{\star})$ may be inverted to describe the pre-contrast relaxation rates:

$$R_{10} = 1/T_{10}; \quad R_{20}^{\star} = 1/T_{20}^{\star} \tag{2.6}$$

For signal-enhancing CAs the resulting relaxation rates can be expressed in terms of the pre-contrast relaxation rates, the relaxivity of the agent $(r_1, r_2^{\star})$ and the concentration of the agent (C).

$$R_1 = R_{10} + r_1 C; \quad R_2^{\star} = R_{20}^{\star} + r_2^{\star} C \tag{2.7}$$

Relaxation rates contribute to the measured DCE-MRI signal. A typical acquisition sequence for DCE-MRI is a spoiled gradient echo (SPGR) with the theoretical equation for signal evolution given by (Mansfield & Morris, 1982; Fram *et al.*, 1987):

$$S(t) = A_0 \sin(\phi) \frac{1 - e^{-T_R R_1(t)}}{1 - \cos(\phi) e^{-T_R R_1(t)}} e^{T_E R_2^{*}(t)} \tag{2.8}$$

Here, A_0 is scaling term dependent on scanner setup and proton density, ϕ is the flip angle, T_R and T_E are the repetition and echo times, respectively. To produce T_1-weighted images, T_E is chosen to be short to suppress T_2^{*} effects. Choice of T_R and ϕ is determined by required sampling interval and signal magnitude.

Assuming a proportional affect of concentration on the signal, $S(t)$, the indicator concentration, $C(t)$, may be defined utilising the pre-contrast relaxation time, T_{10}, and averaged pre-contrast signal, S_0, with r_1, by:

$$C(t) = \frac{1}{r_1 T_{10}} \left(\frac{S(t)}{S_0} - 1 \right), \tag{2.9}$$

as quoted in Wake *et al.* (2018). Here, r_1 is a measured quantity specific to the

CA used, determined experimentally using approaches outlined in Pintaske *et al.* (2006). Alternatively, a non-linear relationship between signal and concentration is derived from equation 2.7 where $R_1(t)$ can be found via solution of equation 2.8.

The T_{10} baseline or T_1 relaxation time prior to the introduction of contrast may be estimated ahead of the main scanning sequence. A variable flip angle approach is usually applied, which acquires two or more SPGR sequences with distinct flip angles.

DCE-MRI data contains the time evolution of tissue CA concentration in the scan volume, which is an average over each measurement voxel and as such does not reflect any internal structures that may be present at a sub-voxel scale. In clinical DCE-MRI research voxel size varies dependent on the type of scanner; in recent work on liver perfusion, voxels from $0.7 \times 0.7 \times 2\text{mm}^3$ to $1.4 \times 1.4 \times 2\text{mm}^3$ have been used (Simeth *et al.*, 2018). Compared to these voxel dimensions many biological features are indistinguishable. This sub-structure may be represented via a compartment model based on the physiology of the tissue type concerned. These approaches are usually referred to as TK modelling and are discussed further in Sections 2.3.3 and 2.4.

2.3.3 Physiological parameters from concentration

To extract useful features, such as blood-brain barrier leakage or CA uptake into tissues, from the DCE-MRI data the resulting concentration-time curves must be analysed further. The analysis methods that can be applied are qualitative, semi-quantitative or fully quantitative. The former two techniques are discussed here as they are straightforward, the latter will be explored in the next section.

Qualitative Methods

For qualitative analysis, a human observer will classify the concentration-time curves based on the initial rise and delayed phase behaviours. These classifications while of use for diagnosing breast cancer and synovial sarcoma (Kuhl *et al.*, 1999; van Rijswijk *et al.*, 2001; Teo *et al.*, 2014), are hindered by high variability in classification between observers due to the subjective nature of the technique.

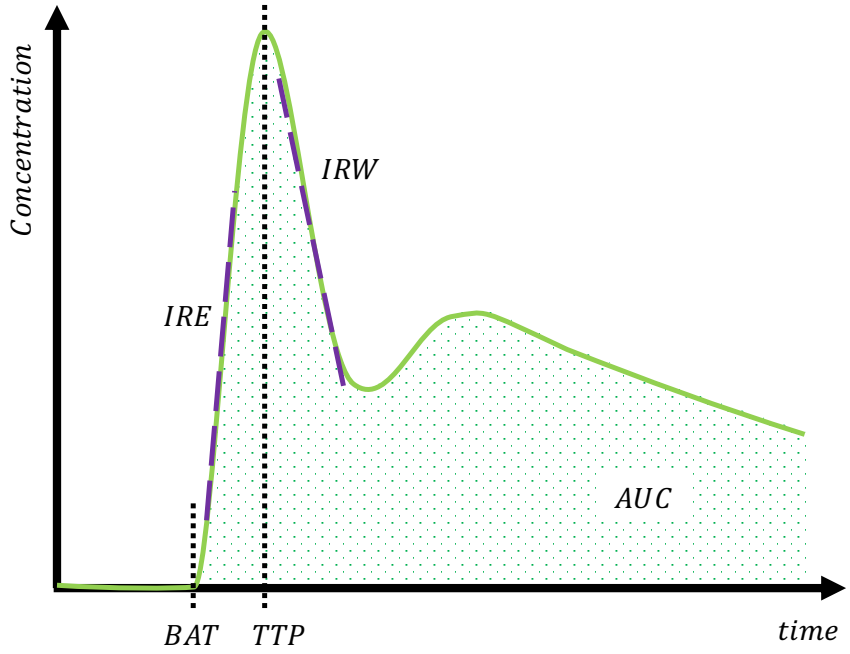

Figure 2.2: Diagram showing examples of various semi-quantitative parameters for DCE-MRI analysis. The concentration curve is denoted by the green line, with the corresponding AUC in green dots. The BAT and TTP are also indicated alongside the initial rates of enhancement and washout (IRE, IRW).

Additionally, no spatial data can be extracted unless shape characterisation is applied on a voxel-wise basis.

Semi-Quantitative Methods

For semi-quantitative analysis the concentration or signal time courses are interrogated to produce numerical values that define certain properties. Some examples of these features are the time-to-peak (TTP), the area under the curve (AUC), the bolus arrival time (BAT) and the mean transit time (MTT) (Cuenod & Balvay, 2013). These features are indicated for a representative signal in Figure 2.2.

A distinct advantage of this semi-quantitative approach is in the simplicity of application to any DCE-MRI data regardless of the system type. By contrast, fully quantitative analysis requires a significant amount of additional work. In this respect, it is of note that quantities such as area under the curve relate to clinically relevant quantitative parameters (Chih-Feng *et al.*, 2012), while being much easier to acquire.

Quantitative Methods

The framework of TK modelling is used to determine physiological quantities in tissues such as blood flow and volume, diffusion and extracellular space volume (Perl *et al.*, 1975; Lassen & Perl, 1979; Koh *et al.*, 2011). It is a technique applied to model the spatial distribution and evolution of CA concentrations. TK approaches model the region of interest (ROI) or tissue as a volume with separate internal compartments. The number of compartments and interactions between them is chosen dependent on the type of tissue being described and the properties of the CA.

In the 1990s, Larsson, Tofts, Brix and their collaborators applied TK principles to extract parameters useful in quantifying blood brain barrier permeability from DCE-MRI results (Larsson *et al.*, 1990; Tofts & Kermode, 1991; Brix *et al.*, 1991). These approaches had a widespread impact on the collection and analysis techniques used for DCE-MRI data.

Advances in DCE-MRI technology have improved temporal resolution and contrast-to-noise ratio of the scan data. This has allowed models to include both Permeability-Surface Area Product (PS) and plasma flow (F_p), which were previously indistinguishable (Sourbron & Buckley, 2011). Since these developments, TK models have been used for analysis in many aspects of DCE-MRI, including motion correction (Flouri, 2016).

2.3.4 Compartmental Modelling

The technique of compartmental modelling uses the premise that a space can be modelled accurately by well mixed interacting sub-spaces. For the case of medical modelling, for example, a space may have venous and arterial contributions. These contributions can be represented as different compartments that themselves are not well mixed but undergo interactions based on their individual concentrations. A summary of processing steps between DCE-MRI data and compartment modelling approaches is detailed in Figure 2.3.

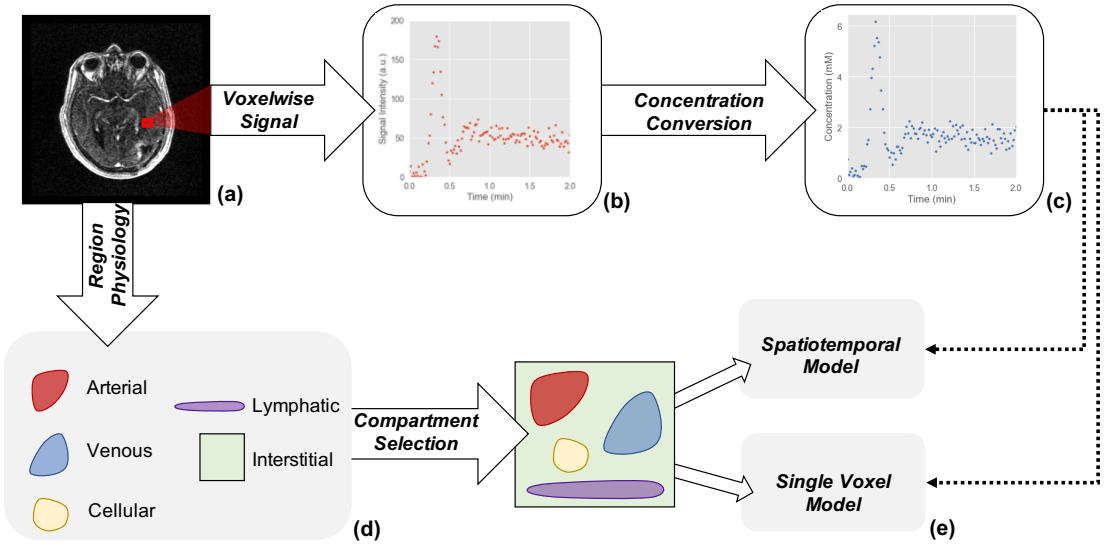

Figure 2.3: Illustrative pipeline for the analysis of DCE-MRI data. Showing an example slice of brain DCE-MRI (Barboriak, 2015) (a) that has a signal extracted across each voxel (b) which can be converted into a CA concentration (c). A model matching the physiology of the region should be selected for fitting (d). The voxelwise concentration-time data can then be analysed using an appropriate spatiotemporal or single voxel approach (e).

Local and Tissue Concentrations

As DCE-MRI data is only acquired at the voxel level, the tissue concentration for each compartment is calculated from the distributed volume fractions, v, and summed to give the average tissue concentration of the voxel, given generally by (Sourbron, 2014):

$$C_i(t) = \sum_{q=1}^{Q} C_i^q(t) = \sum_{q=1}^{Q} v^q c^q(t). \tag{2.10}$$

In (2.10) there are a total number of Q compartments in voxel i, with tissue, C, and local, c, concentrations.The local compartment concentration, c^q, is the concentration within compartment q, the tissue compartment concentration, C^q, is the concentration of compartment q spread over the voxel volume according to the compartmental volume fraction. Descriptions may be detailed in terms of either c or C, which will be referred to as the local or tissue concentration picture, respectively.

Two Compartment Exchange Model

The two compartment exchange model (2CXM) equations (2.11-2.12) describe the evolution of concentration in the plasma, c^p, and extracellular volume, c^e, compartments:

$$v^p \frac{dc^p(t)}{dt} = F_p c_a(t) + PS c^e(t) - (F_p + PS) c^p(t) \tag{2.11}$$

$$v^e \frac{dc^e(t)}{dt} = PS(c^p(t) - c^e(t)) \tag{2.12}$$

The arterial input function (AIF) is represented by $c_a(t)$ in (2.11). This input concentration-time curve is usually a measured quantity from a large upstream artery. A comprehensive review of use and measurement of the AIF can be found in Calamante (2013). When measurement of the AIF is not possible a high-resolution population averaged AIF can be applied (Parker *et al.*, 2006). Alternative work towards an inferred AIF from measurements is shown in Jiřík *et al.* (2019, 2022). At present, the analysis based on the compartmental model makes use of the AIF, assuming every voxel in the ROI receives this inflow. Consequently, this neglects local exchange between and within voxels in the ROI.

Tofts Models

One of the most widely applied compartmental models, the Tofts model (Tofts, 1997), includes one tissue compartment as well as a vascular compartment. This is a simplification of a general 2CXM, with the plasma volume, v_p, taken as negligible. The model is described as:

$$C(t) = K^{trans} e^{-t\frac{K^{trans}}{v_e}} * c_a(t) \tag{2.13}$$

This has since had more complexity introduced to form the Extended-Tofts model (Tofts *et al.*, 1999), by allowing an intravascular contribution, by inclusion of v_p:

$$C(t) = v_p c_a(t) + K^{trans} e^{-t\frac{K^{trans}}{v_e}} * c_a(t) \tag{2.14}$$

The extension of the model allowed more highly vascularised physiology to be modelled, such as cancerous tissues. A comprehensive overview of different TK models, the associated parameters and physical interpretations can be found in Sourbron & Buckley (2012, 2013).

Here K^{trans} represents the volume transfer constant defining the leakage of indicator from the plasma into the extracellular extravascular space. These models are only accurate when applied to weakly vascularised (Tofts and Extended-Tofts) or highly perfused tissues (Extended-Tofts) (Sourbron & Buckley, 2011).

In a non-MRI context the Kety (1951) model was proposed for measurement of blood flow via the exchange of inert gas within the body. The Kety model is mathematically equivalent to the Tofts model, as such the general approach in the literature is referred to as the Tofts-Kety, Tofts or Kety model. In this thesis, the Kety model is only discussed in Chapter 4 due to its specific use within one of the articles considered.

Implementation for Model Fitting

There are two main approaches for the implementation of compartmental modelling in order to return physiological parameters. Firstly, approaches may apply non-linear least squares methods which fit the model outcome to the measurement data. These approaches typically apply a gradient descent-based method that sequentially update initial guesses for parameter values. Secondly, approaches may apply linear least squares methods which solve for parameters based on a system of linear equations. These linear methods are increasing in popularity due to minimised computational time and no required initial guesses. Applications of linear methods have been illustrated for the Tofts models (Murase, 2004) and the 2XCM by Flouri *et al.* (2016).

2.4 Spatiotemporal Tracer Kinetics

Many standard methods apply the *central volume theorem* (Perl *et al.*, 1975) which assumes the accessible tissue volume can be calculated using the incoming

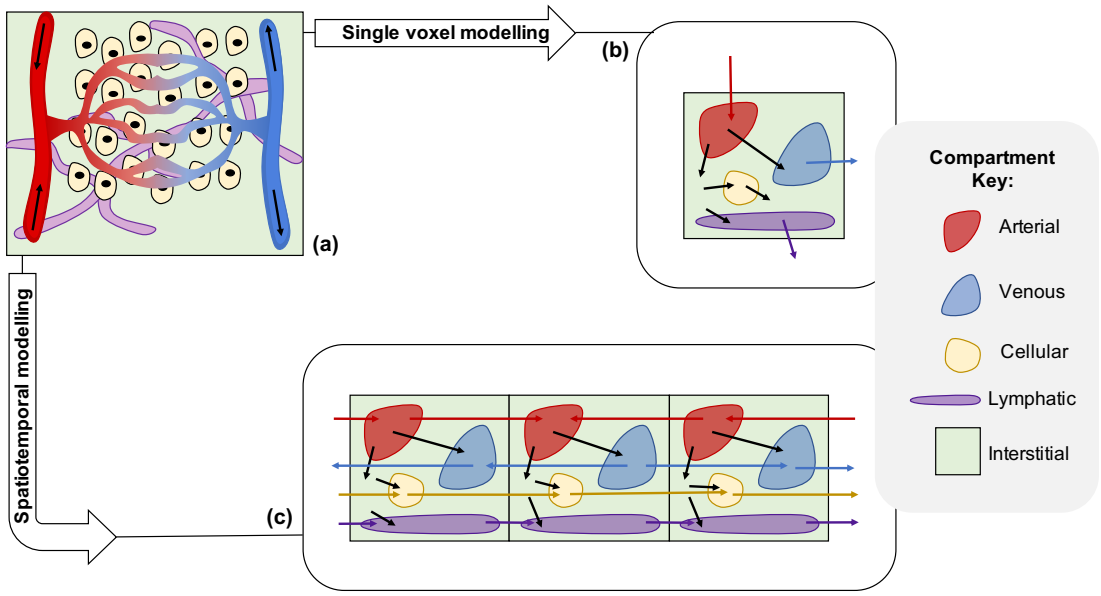

Figure 2.4: Illustration of potential tissue physiology including arterial, venous, lymph tissue and interstitial spaces (a) that may be modelled via compartments in either single voxel (b) or spatiotemporal approaches (c).

flow and MTT. This theorem is only valid for linear stationary tissues with a singular inlet where transport proceeds in a purely convective manner. Systems with multiple inlets which have no equivalent single inlet, and either diffusive or asymmetric transport (such as the kidney (Sourbron & Buckley, 2012)) cannot be described by this approach. To model transport parameters with a high spatial resolution, voxels should be treated as interacting spaces (i.e. multiple inlets) which could have non-convective transport meaning methods applying the *central volume theorem* are no longer appropriate.

The mathematical framework applied throughout the majority of work within this thesis adheres to the theoretical basis set forward for spatiotemporal TKs by Sourbron (2014). A notable and useful feature of this framework is the generality of the descriptions which allow flexibility to define any system that may be of interest. The following sections detail the key principles of spatiotemporal TK theory as defined by Sourbron (2014). An illustration of single voxel and spatiotemporal compartment modelling approaches is detailed in Figure 2.4.

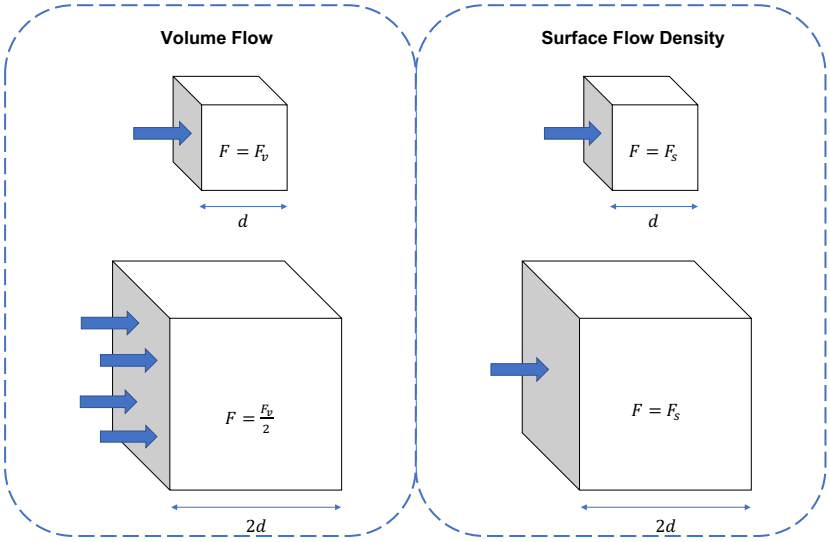

Figure 2.5: Illustration of the scaling problems with definition of perfusion from volume based flow (left) and the removal of this for surface flow density approaches (right).

2.4.1 Transport Processes

CA is transported throughout the body after injection into the vascular system. Within the body there are many different physiologies with different transport processes controlling the evolution of concentration. In the vascular system transport is convection dominated and diffusion has a negligible impact on the transport. Conversely, in the interstitial space convective transport is reduced and both diffusive and convective transport can play a role. Transport by diffusion or active transport can occur across cell membranes. Within the lymphatic system, skeletal muscle and vessel wall contractions produce convective flow in collecting vessels, while in lymph nodes both convection and diffusion processes are relevant (Don *et al.*, 2024; Thorup *et al.*, 2023).

With advective transport, appropriate physical interpretations of the blood flow must be put forward. Definition of perfusion as the volume per unit time into a voxel, as shown in Figure 2.5 (left), produces values that depend on the voxel volume itself. Henkelman (1990) highlighted this as a problematic definition. Conversely, if the surface flow density, or volume per unit time per unit area, is considered perfusion can be defined as independent of the voxel size (Figure 2.5 (right)), as in Sourbron (2014). Surface flow densities are applied within the spatiotemporal models defined below.

The advective flux, $\vec{j}$, in the local and tissue concentration picture is represented in (2.15) and (2.16), respectively.

$$\vec{j}(\vec{r},t) = \vec{f}(\vec{r})c(\vec{r},t) \tag{2.15}$$

$$\vec{j}(\vec{r},t) = u^{\star}(\vec{r})C(\vec{r},t) \tag{2.16}$$

Where $\vec{f}$ represents the surface flow density, and $u^{*} = \frac{\vec{f}}{v}$ the blood velocity. The diffusive flux is shown by equations 2.17 and 2.18. Diffusion can be described in both the local and tissue concentration referred to as molecular, $D^{\star}$, or pseudo, D, diffusion, respectively.

$$\vec{j}(\vec{r},t) = -D^{\star}(\vec{r})\vec{\nabla}c(\vec{r},t) \tag{2.17}$$

$$\vec{j}(\vec{r},t) = -D(\vec{r})\vec{\nabla}C(\vec{r},t) \tag{2.18}$$

Here c and C are related as defined in (2.10). The conversion shown in equations 2.17 and 2.18 is only valid for systems with a constant v, which may be moved through the gradient operator, so the pseudo diffusion is defined by (2.19).

$$D(\vec{r}) = \frac{D^{\star}(\vec{r})}{v} \tag{2.19}$$

2.4.2 Consequence of Representation

While the local concentration picture is more intuitive as v is explicitly included, the tissue concentration picture is more useful for numerical description of the system. For example, when considering C it is unnecessary to track the flow incompressibility (Sourbron, 2014).

As the diffusive flux is dependent on the spatial gradient of the concentration, transforming between the local and tissue concentration pictures has consequences when the volume fraction is not constant and $v = v(\vec{r})$. For heterogeneous $v(\vec{r})$ it is no longer valid to extract as a constant through the gradient operator as in (2.18), the correct relation to be considered is shown by

(2.20).

$$\vec{j}(\vec{r},t) = -D^{\star}(\vec{r})\vec{\nabla}\left(\frac{C(\vec{r},t)}{v(\vec{r})}\right) \tag{2.20}$$

Expansion of (2.20) gives (2.21), where, due to the spatial gradient in the volume fraction, a pseudo advection term is created on application of the gradient operator. The direction of this term follows the gradient of $v(\vec{r})$.

$$\vec{j}(\vec{r},t) = -D^{\star}(\vec{r})C(\vec{r},t)\vec{\nabla}\left(\frac{1}{v(\vec{r})}\right) - \frac{D^{\star}(\vec{r})}{v(\vec{r})}\vec{\nabla}C(\vec{r},t) \tag{2.21}$$

To account for these compound effects generally the total flux may be written as:

$$\vec{j}(\vec{r},t) = \vec{u}(\vec{r})C(\vec{r},t) - D(\vec{r})\vec{\nabla}C(\vec{r},t). \tag{2.22}$$

With the coefficients u and D described by the relations in (2.23).

$$\vec{u}(\vec{r}) = \frac{\vec{f}(\vec{r})}{v(\vec{r})} + \frac{D^{\star}(\vec{r})}{v(\vec{r})^2}\vec{\nabla}v(\vec{r}) \quad ; \quad D(\vec{r}) = \frac{D^{\star}(\vec{r})}{v(\vec{r})} \tag{2.23}$$

This picture collapses to the constant uniform volume picture for isotropic fractions as the gradient and resulting pseudo advection term vanishes. Only the pseudo diffusion term associated with moving from the local into the tissue concentration representation remains.

2.4.3 System Examples

Following the above discussion it is straightforward to produce spatiotemporal TK equations which capture systems with any number of compartments or transport processes present. For the subject matter of this thesis, focus is given to one-compartment vascular systems and two-compartment perfusion systems. Spatiotemporal equations governing each system type are described within this section, giving details of systems with either diffusive and advective transport, or solely advective transport. One-compartment systems considering solely diffusive transport can be used for appropriate physiology such as regions of necrosis.

One-Compartment Vascular System

For singular or one-compartment systems description is rather simple. For description in the local concentration picture this results in:

$$v(\vec{r})\frac{\partial c(\vec{r},t)}{\partial t} = -\vec{\nabla}\cdot\vec{f}(\vec{r})c(\vec{r},t) + \vec{\nabla}\cdot D^*(\vec{r})\vec{\nabla}c(\vec{r},t) \tag{2.24}$$

Within vascular systems of negligible diffusion, all definitions within sections 2.4.1 and 2.4.2 are adjusted by setting $D = 0$. Therefore, without diffusive transport, equation 2.24 would be reduced to:

$$v(\vec{r})\frac{\partial c(\vec{r},t)}{\partial t} = -\vec{\nabla}\cdot\vec{f}(\vec{r})c(\vec{r},t) \tag{2.25}$$

For the tissue concentration picture the conversion yields these to give:

$$\frac{\partial C(\vec{r},t)}{\partial t} = -\vec{\nabla}\cdot\vec{u}(\vec{r})C(\vec{r},t) + \vec{\nabla}\cdot D(\vec{r})\vec{\nabla}C(\vec{r},t) \tag{2.26}$$

$$\frac{\partial C(\vec{r},t)}{\partial t} = -\vec{\nabla}\cdot\vec{u}(\vec{r})C(\vec{r},t) \tag{2.27}$$

Additionally, total flow in these listed one-compartment systems is assumed incompressible, *i.e.* $\vec{\nabla}\cdot\vec{f} = 0$. Therefore, 3 degrees of freedom are illustrated for the local concentration picture: with 4 scalar fields $(v, \vec{f})$, and one degree of freedom removed by flow incompressibility. Similarly, for the tissue concentration picture there are 3 degrees of freedom without constraint as only the velocity with 3 scalar fields $(\vec{u})$ is applied (Sourbron, 2014).

Two-Compartment Perfusion System

For a description of two-compartment systems, there is a need for additional transport processes to define interactions which exist between each of the two compartments within the system. Within this thesis, spatiotemporal two-compartment systems are specifically limited to arterio-venous systems with mono-directional exchange mediated via perfusion. This perfusion definition emerges naturally within the theoretical setup (Sourbron, 2014) as the

divergence of the arterial flow.

$$F = -\vec{\nabla} \cdot \vec{f^a} \tag{2.28}$$

Here, as in the one-compartment systems, the total flow remains incompressible, *i.e.* $\vec{\nabla} \cdot (\vec{f^a} + \vec{f^v}) = 0$.

Explicitly these two-compartment systems are described for the local concentration picture as:

$$v^a(\vec{r})\frac{\partial c^a(\vec{r},t)}{\partial t} = -\vec{\nabla} \cdot \vec{f^a}(\vec{r})c^a(\vec{r},t) + \vec{\nabla} \cdot D^{a*}(\vec{r})\vec{\nabla}c^a(\vec{r},t) - Fc^a \tag{2.29}$$

$$v^v(\vec{r})\frac{\partial c^v(\vec{r},t)}{\partial t} = -\vec{\nabla} \cdot \vec{f^v}(\vec{r})c^v(\vec{r},t) + \vec{\nabla} \cdot D^{v*}(\vec{r})\vec{\nabla}c^v(\vec{r},t) + Fc^a \tag{2.30}$$

With the corresponding tissue concentration equations given straightforwardly, using the additional relation $F = K^{va}v^a$, by:

$$\frac{\partial C^a(\vec{r},t)}{\partial t} = -\vec{\nabla} \cdot \vec{u}^a(\vec{r})C^a(\vec{r},t) + \vec{\nabla} \cdot D^a(\vec{r})\vec{\nabla}C^a(\vec{r},t) - K^{va}C^a \tag{2.31}$$

$$\frac{\partial C^v(\vec{r},t)}{\partial t} = -\vec{\nabla} \cdot \vec{u}^v(\vec{r})C^v(\vec{r},t) + \vec{\nabla} \cdot D^v(\vec{r})\vec{\nabla}C^v(\vec{r},t) + K^{va}C^a \tag{2.32}$$

For systems with negligible diffusion these can be condensed in the same way as the one-compartment systems, to give the local concentration relations in Equations 2.33-2.34.

$$v^a(\vec{r})\frac{\partial c^a(\vec{r},t)}{\partial t} = -\vec{\nabla} \cdot \vec{f^a}(\vec{r})c^a(\vec{r},t) - Fc^a \tag{2.33}$$

$$v^v(\vec{r})\frac{\partial c^v(\vec{r},t)}{\partial t} = -\vec{\nabla} \cdot \vec{f^v}(\vec{r})c^v(\vec{r},t) + Fc^a \tag{2.34}$$

With the corresponding tissue concentration equations given straightforwardly, using the additional relation $F = K^{va}v^a$, by:

$$\frac{\partial C^a(\vec{r},t)}{\partial t} = -\vec{\nabla} \cdot \vec{u}^a(\vec{r})C^a(\vec{r},t) - K^{va}C^a \tag{2.35}$$

$$\frac{\partial C^v(\vec{r},t)}{\partial t} = -\vec{\nabla} \cdot \vec{u}^v(\vec{r})C^v(\vec{r},t) + K^{va}C^a \tag{2.36}$$

2.5 Physics Informed Neural Networks

Within this thesis, Chapter 6 concerns the application of PINNs for the retrieval of parameters in spatiotemporal systems. This section details some supporting background material to aid the reader in defining terms and ideas generally found in many machine learning approaches and specifically within standard and more advanced neural networks (NNs) of which PINNs are an example.

2.5.1 Neural Networks

Standard NNs have a structure consisting of an input layer, some hidden layers and an output layer: an overview of a typical setup is shown in Figure 2.6. The inputs usually take the form of system coordinates such as spatial or temporal points (x, y, z, t). The hidden layers are the bulk of the network that pass the values from the input layer through the network, via weighted connections and activation functions that act at each node within the network (Mehlig, 2021). At the final layer, the output is exposed as an array of values resulting from the effect of the network on the input values. The outputs typically represent a measurable quantity, or a quantity that may be derived from available data. A cost function is defined using the output layer values and the known measurement or simulation values, $\mathcal{L}_{Data}$. This cost function is then minimised by carrying out a backpropagation step across the network in order to update the weights to reduce the defined cost function (Brunton & Kutz, 2019, Chp. 6).

In this way, the neural network can act as a complex operator that 'learns' to transform input coordinates into the observed quantity (Cybenko, 1989). These network types are highly reliant on the training data that is used during the initial phase where weights are updated. As such, sparse or unrepresentative training sets can bias the network and reduce the robustness of outputs. Additionally, in some cases the network will not converge.

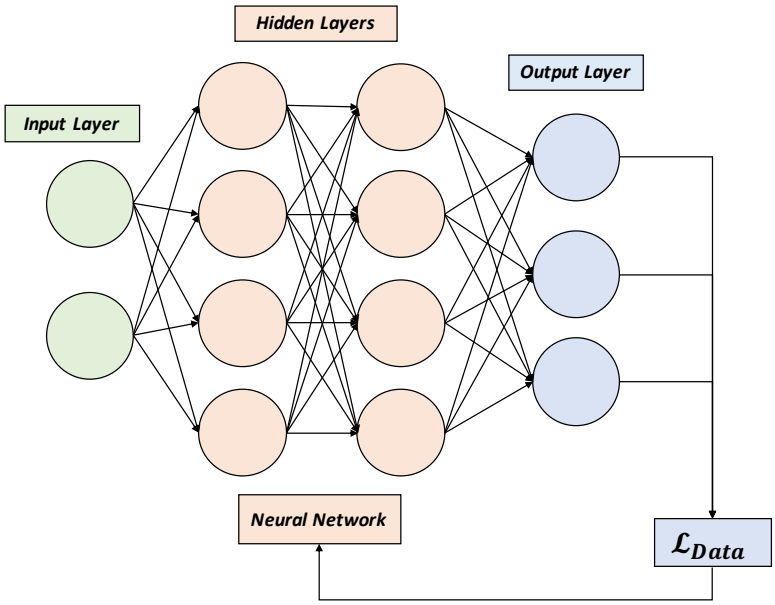

Figure 2.6: Representation of a standard neural network.

2.5.2 Automatic Differentiation

All differentiation within a PINN is carried via automatic differentiation (Baydin *et al.*, 2018). This relatively 'new' approach has been around for about 40 years (Rall, 1981) but has gained attention through its application in the machine learning field (Griewank & Walther, 2008). It is a clever numerical technique which enables exact derivatives to be calculated and differs from symbolic or numerical differentiation.

Within automatic differentiation, a dual number, N_{dual} is defined.

$$N_{dual} = x + \epsilon \tag{2.37}$$

Where ϵ is an abstract number with the defined property $\epsilon^2 = 0$. Expanding a function $f(x)$ about ϵ using a typical Taylor's series results in a familiar equation.

$$f(N_{dual}) = f(x+\epsilon) = f(x) + \epsilon\frac{\mathrm{d}f}{\mathrm{d}x} + \frac{1}{2!}\epsilon^2\frac{\mathrm{d}^2 f}{\mathrm{d}x^2} + \frac{1}{3!}\epsilon^3\frac{\mathrm{d}^3 f}{\mathrm{d}x^3} + ... \tag{2.38}$$

Due to our previously defined behaviour of ϵ all terms after the first two are automatically equal to zero and are therefore eliminated. Following the outcome of a dual number for the expansion in a typical function, we see the derivative

term emerge with ϵ as a coefficient (Baydin *et al.*, 2018).

$$f(x) = x^2 + 2x \tag{2.39}$$

$$f(x+\epsilon) = (x+\epsilon)^2 + 2(x+\epsilon) \tag{2.40}$$

$$f(x+\epsilon) = x^2 + 2x + \epsilon(2x+2) = f(x) + \epsilon\frac{\mathrm{d}f}{\mathrm{d}x} \tag{2.41}$$

$$\therefore \frac{\mathrm{d}f}{\mathrm{d}x} = 2x + 2 \tag{2.42}$$

This approach is applied throughout a PINN for calculating derivative terms with respect to input coordinates and also for back-propagating the effect of the weight at each neuron on the overall loss function.

2.5.3 Activation Functions

Each node in the network takes incoming values from every other node in the previous layer. Each connection has its own trainable weight that adjusts the value incoming to the node. The summation of all these inputs at a node is passed to an activation function to then be transmitted to the next layer (Mehlig, 2021). There are several different types of activation functions that have been employed during the advent of standard NNs and PINNs. Activation functions are a way to introduce non-linearity into the network: without them all hidden layers collapse as the final layer is just a linear function of the first layer thus discounting any additional layer found in between (Brunton & Kutz, 2022, Chp. 14). Indeed, the non-linearity of activation functions has been proved to allow NNs to function as universal function approximations (Cybenko, 1989).

Binary Activation Functions

Binary activation functions act by returning -1 until a threshold value is reached and then returning 1, a classic step function (McCulloch & Pitts, 1943). The problem in the network comes from the poorly defined gradient segment at the step change, and the flat gradients at all other points rendering this a poor choice for activation functions (Snyman, 2005).

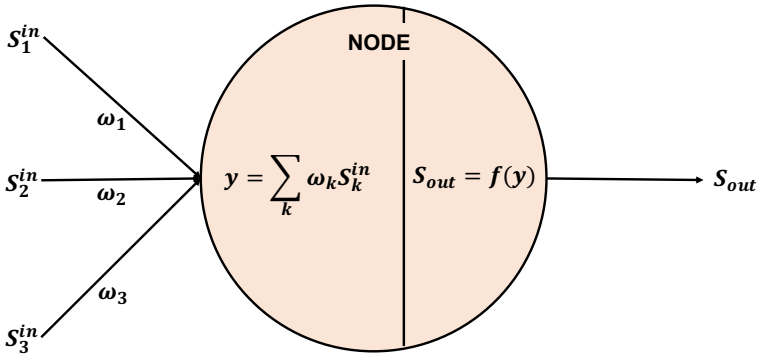

Figure 2.7: Representation of a node in any type of Neural Network, showing incoming values on neurons, S_k^{in}, with distinct weightings, ω_k. These are combined via a weighted summation y, which is then converted to the outgoing value from the node, S_{out}, using the activation function $f(y)$.

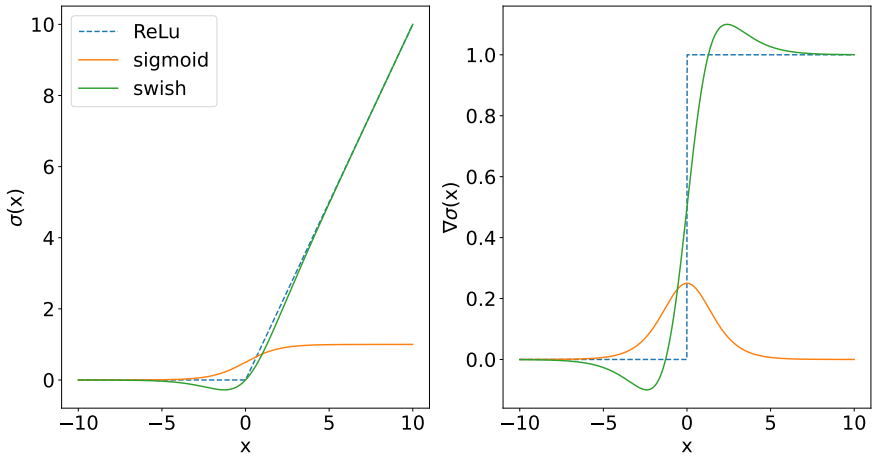

Figure 2.8: Common activation functions showing values for functions (left) and first derivative values (right).

Continuous Activation Functions

Many different types of activation functions that have been developed over the years (Figure 2.8). The popular choices have been the sigmoid, ReLU (Nair & Hinton, 2010) and tanh activation functions. Recently, the swish function, as proposed by researchers at Google (Ramachandran *et al.*, 2017), has shown promise for fixing the vanishing gradient problem.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{2.43}$$

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{2.44}$$

$$\text{ReLU}(x) = \max(0, x) \tag{2.45}$$

$$\text{swish}(x) = \frac{x}{1 + e^{-x}} \tag{2.46}$$

2.5.4 Addition of Physical Laws

A significant drawback with classical machine learning techniques is the way the network arrives at the optimum, as the underlying physics of a system is not taken into account, with only the overall cost function being reduced. This has the potential to return unphysical solutions if outputs are not fully interrogated. In many systems, the underlying physics or governing equations could be a key component in arriving at the optimum physically acceptable solution.

2.5.5 PINNs Development

Since the introduction of PINNs in the now seminal paper from Raissi *et al.* (2019), interest in applications for this new technique has been widespread and numerous. Indeed at the time of writing the original manuscript has gathered over 5,800 citations and counting. Ideas put forward in this paper collect work from their previous two-part preprint (Raissi *et al.*, 2017a,b), which details not only a method for the solution of forward modelling but also the inverse problem of parameter recovery in non-linear systems.

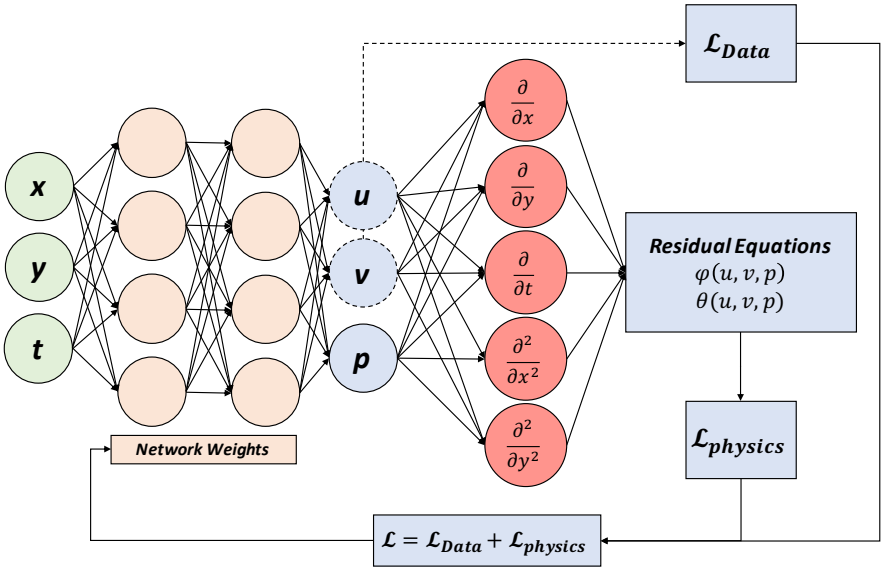

Figure 2.9: Representation of a PINN architecture.

2.5.6 PINNs Implementaion

From an implementation perspective, a PINN very closely resembles a standard NN, except with additional operations on the output layer (Raissi *et al.*, 2019; Cuomo *et al.*, 2022). Here any necessary derivatives are taken of parameters in the output layer with respect to the physical coordinates describing the system. These terms are then used to define residual equations that fully characterise the underlying physics of the system. Equations must be cast in residual form so that minimisation is straightforward. There may be any number of residual equations that can describe not only the overall physics of the system but also any known initial and boundary conditions.

A summation of these residual terms is then carried out to define the physics loss function $\mathcal{L}_{Physics}$:

$$\mathcal{L}_{Physics} = \sum_i \phi_i \tag{2.47}$$

This additional loss term is then combined with $\mathcal{L}_{Data}$ to give a new overall loss function for the system:

$$\mathcal{L} = \mathcal{L}_{Data} + \mathcal{L}_{Physics} \tag{2.48}$$

2.6 Summary

This chapter has provided a summary of background material relevant for the research presented within this thesis. Specifically, an overview of DCE-MRI, single voxel analysis techniques of DCE-MRI data, spatiotemporal theory for DCE-MRI analysis and principles underlying PINNs approaches have been covered.

Chapter 3

The OSIPI-DCE challenge

Pharmacokinetic (PK) analysis techniques have historically been developed as in-house research codes with implementation differences between studies. These methods are often developed with and applied to clinical datasets which are not freely accessible due to data protection concerns. The lack of widely accessible clinical datasets has hindered the comparison of PK modelling outcomes (Carp, 2012). Differences in reported parameters from differing implementation methods (O'Connor *et al.*, 2007) require further characterisation to ensure quantitative evaluation of DCE-MRI can yield reproducible results. Initiatives such as the Cancer Imaging Archive (TCIA) (Clark *et al.*, 2013) are invaluable for benchmarking PK modelling by allowing different research groups free access to de-identified datasets.

This chapter[1] is comprised of a detailed investigation of 10 separate software packages to measure and characterise differences in the K^{trans} parameter when applied to the same DCE-MRI datasets. Metrics based on accuracy, repeatability and reproducibility are utilised in order to provide quantitative insight into differences which appear in the resulting K^{trans} values. The associated scoring and synthetic data production codes to the material presented can be accessed via Github[2].

[1]The work in this chapter has been published (Shalom *et al.*, 2024b), this is defined in 'Joint Publications' with contribution details on Page ii. Some Figures (3.5,3.6,3.9,3.10,3.7) and Tables (3.4) from the original Supporting Information are included in the main text. Large data tables from the original Supporting Information are included in Appendix A.

[2]https://github.com/OSIPI/TF6.2_DCE-DSC-MRI_Challenges

3.1 Introduction

DCE-MRI provides physiological parameters associated with the exchange of a CA between intravascular space and extravascular spaces (Bell *et al.*, 2021). In patients with glioblastoma, K^{trans} has been proposed as a marker for characterising tumour pathophysiology, which can aid in grading (Zhang *et al.*, 2012), assessment of treatment response (Kickingereder *et al.*, 2015), and differentiation of recurrence from radiation necrosis (Thomas *et al.*, 2015).

Quantitative DCE-MRI through PK modelling is intended to yield reproducible parameters across different studies and institutions (Zhu *et al.*, 2005). However, the variation in the AIF, chosen PK models, and stability in model fitting adversely affect the quantification of K^{trans} values (O'Connor *et al.*, 2007). Therefore, the reported K^{trans} values differ among studies, making it currently unsuitable as a marker in multi-institutional clinical trials. Furthermore, a small number of studies have measured repeatability. Based on this limited literature, current QIBA guidelines state that a change of K^{trans} above 21.3% may indicate true K^{trans} change with 95% confidence in glioblastoma (QIBA MR Biomarker Committee, 2020; Shukla-Dave *et al.*, 2019), if no repeatability data is available for the study. Therefore, methods with a repeatability coefficient (%RC) (Shukla-Dave *et al.*, 2019) above this threshold cannot reliably detect tumour progression in longitudinal studies, further contributing to the limitations of quantitative DCE-MRI in clinical trials.

Over the past decade, more attention has been brought to the replication of research studies, the so-called reproducible research (Lindquist, 2020; Stikov *et al.*, 2019). While researchers make their best effort to report accurate data, the choices they have to make about different aspects of data collection and analysis methods could influence the outcome of their significance tests and therefore, the derived conclusions (Simmons *et al.*, 2011; Wicherts *et al.*, 2016). This "researcher degrees of freedom" issue imposes challenges for the reproducibility of the results when reanalysing the same data, or generalisability of the findings to independent data (Simmons *et al.*, 2011; Wicherts *et al.*, 2016).

3. The OSIPI-DCE challenge

For quantification of K^{trans} from DCE-MRI, there is an extensive list of available tools from which to choose although no "gold standard" analysis technique exists for clinical data. Evaluation and validation of these tools in the reported literature are based on datasets collected by authors, rendering it difficult to perform a fair comparison between them (Carp, 2012). When provided with a wide range of possible (well-grounded) options for analysis methods, researchers may select or report methods that yield more favourable results for their data (Simmons *et al.*, 2011; Wicherts *et al.*, 2016). Despite researchers' best intentions, the inclination to show statistically significant results, referred to as "selective analysis reporting", could accompany false positive errors (Carp, 2012). To avoid such errors, which hinder reproducibility, it is critical to provide comprehensive and open/transparent details about the study design and analysis approaches. The International Society for Magnetic Resonance in Medicine Open-Science Initiative for Perfusion Imaging (ISMRM-OSIPI), referred to hereafter as "OSIPI"), an ISMRM perfusion study group initiative, was founded to promote reproducible research and open science in perfusion imaging and to facilitate the translation of software tools into clinical practice. The OSIPI task force on DCE and DSC challenges (Task Force 6.2), was formed in February 2020 with a group of medical physicists, radiologists, and biomedical engineers to address the current issues of benchmarking perfusion quantification methods by organising community challenges. The OSIPI-DCE, as an ISMRM challenge, was the first of such challenges. OSIPI- DCE aims to design and build a systematic and controlled framework to benchmark the quantification of K^{trans} as a diagnostic or prognostic biomarker in glioblastomas, and to apply this framework to submissions from the community. This setup allowed the evaluation and validation of software packages in a single setting with synthetic and real-world clinical data. For the first time in a challenge setting, accuracy, repeatability, and reproducibility of various methods were assessed for K^{trans} quantification in brain tumours. This study describes the challenge data, design, results of evaluating different analysis methods, and obstacles in the assessment of reproducibility.

3.2 Methods

The OSIPI-DCE challenge was launched at the ISMRM Annual Meeting, May 15, 2021, upon presentation of the abstract on the outline of the challenge during this annual meeting.

3.2.1 Challenge Setup

This challenge aimed to assess the results and analysis methods submitted by the participating teams according to the $OSIPI_{gold}$ score (Table 3.1) for their 1) accuracy in the quantification of K^{trans} using a set of synthetic data designed for this challenge, 2) repeatability using open access test-retest scans of eight patients with glioblastoma (Clark *et al.*, 2013; Barboriak, 2015), and 3) reproducibility based on an independent re-analysis of the data by a neutral evaluator team.

The challenge design was submitted as an abstract (Kazerooni *et al.*, 2021a) for peer-review at the 2020 ISMRM Annual Meeting. The researchers in the perfusion MRI community were invited by email to participate through the ISMRM Perfusion Study Group, LinkedIn, Twitter, or via direct contact. Interested teams registered on the ISMRM Challenges website and received submission guidelines via an automated email. The participants were asked to submit their results along with a report about the analysis approach, as described below:

- Matrices of voxelwise K^{trans} maps (in the original space) for all slices in the synthetic and clinical DCE-MRI (2 visits per subject) in NIfTI format.
- Standard Operating Procedures (SOPs) with sufficient detail to allow a neutral evaluator team to reproduce the results without interaction with the challenge participants. The SOPs should explain software access and installation and provide a step-by-step guide to reproduce the analysis. It is essential that the synthetic and patient data are analysed with the same approach. Copies of each submission SOP are contained within the OSIPI DCE-Challenge Github.[1]

[1] https://github.com/OSIPI/TF6.2_DCE-DSC-MRI_Challenges/tree/main/ChallengeTeamData.

No requirement was placed on the challenge participants to release their source codes or to base their submissions on open-source or open-access software. However, for commercial or in-house software that was not freely available, the participants were asked to provide a trial license or an executable file for the independent replication of the results. The license could be temporary, allowing sufficient time for re-analysis. Instructions on how to obtain the license should have been included in the SOPs without requiring interactions between submitters and neutral evaluator teams.

The challenge was open for submissions through to the end of the year 2021. The task force reached out to experienced DCE-scientists to help in evaluating the submissions in terms of procedural reproducibility and the reported K^{trans} maps, after the challenge was closed. The evaluators had either more than one (Ivan A. Wolansky), five (Zaki Ahmed, Samuel Bobholz) or ten (Peter S. LaViolette, John Virostko, Hendrik O. Laue, Kyunghyun Sung) years of experience in DCE analysis. At the end of the challenge, the SOPs and software tools for each submission were provided to the independent evaluators.

3.2.2 Data Description

Two sets of data were provided in our challenge repository (Kazerooni *et al.*, 2021b).

Clinical Data

A set of repeat DCE-MRI and T_1-mapping scans with accompanying T_1 CE-FLASH and T_2 CE-FLAIR from 8 patients with glioblastoma, selected from RIDER Neuro MRI database (Clark *et al.*, 2013; Barboriak, 2015) and renamed, acquired on a 1.5T Siemens scanner at two scan dates, typically 1-2 days apart. Sequence details are provided in Appendix A.

Synthetic Data

Two synthetic DCE-MRI patient datasets were generated from RIDER subjects (Clark *et al.*, 2013; Barboriak, 2015) to be analysed with the same processing

pipeline as the clinical data. For this reason, the synthetic DCE-MRI data were integrated into an original DICOM study, including also the anatomical reference data from the same RIDER subjects.

Synthetic data were created following two steps: first, an inverse model was applied to the DCE-MRI and variable flip angle (VFA) data set to obtain parameter maps for pre-contrast relaxation rate (R_{10}), rate constants (K^{trans}, k_{ep}), capillary plasma volume per unit volume of tissue (v_p), and an AIF; subsequently these parameter maps had thresholds and filters applied to produce an unknown ground truth. The forward model was applied to produce synthetic DCE-MRI and VFA signal intensity curves. All details of the inverse modelling remained undisclosed during the submission period. The challenge guideline detailed the PK model used for the forward modelling, as well as the assumed concentration and relaxation rate relationships. Also, it defined the creation of VFA data using R_{10} maps and a constant R^*_{20} throughout.

For the inverse approach, initial parameter values were recovered from the RIDER data using matrix form (Murase, 2004). A partial volume correction was applied using the sagittal sinus signal. Thresholds were then applied to the output to discard negative values produced during the least squares fitting process and limit maximal volume fraction values to 1. Smoothing of the fitted values was carried out using a 3x3 median filter. The AIF signal (corrected with hematocrit 0.45 (Brix *et al.*, 2004)) was selected from the middle cerebral arteries and scaled to have a realistic peak value of 6 mM (Parker *et al.*, 2006).

For the forward model, the extended Tofts model was applied with the AIF and parameter maps (K^{trans}, k_{ep}, v_p). The resulting tissue concentration-time curves were converted into $R_1(t)(= 1/T_1(t))$ and $R^*_{20}(t(= 1/T^*_2(t)))$, assuming a linear relationship between concentration and the relaxation rates according to the r_1 and r^*_2 relaxivities of Gd-DTPA respectively (3.9 and 10 Hz/mM (Pintaske *et al.*, 2006; Siemonsen *et al.*, 2008)), and by making use of the R_{10} and R^*_{20} maps. To deduce the precontrast relaxation rates, the signal evolution was modelled as a SPGR sequence in steady state. This was applied to express the signal everywhere at the initial time and the initial sagittal sinus signal. These relations were combined in order to give a calculated R_{10} map using the reference T_1 of

1.48 s in the sagittal sinus at 1.5T (Zhang *et al.*, 2013). The constant precontrast R^*_{20} applied was 17.24 Hz (Siemonsen *et al.*, 2008). Scan-specific constants match the original scan values with FA= 25°, TR= 3.8 ms, TE= 1.8 ms; using a 1x1x5 mm^3 voxel size.

Subsequently, these relaxation rates were converted into DCE-MRI signal time curves again modeling signal evolution as an SPGR sequence. A multiplicative constant was defined to give the synthetic data similar maximal signal values to the original RIDER dataset. This produced synthetic data with a 16-slice volume captured with a temporal resolution of 4.8 s.

Finally, Rician noise was added to the signal-time data by assessing the standard deviation across the pre-contrast time steps within each voxel from the original RIDER dataset. Voxelwise noise values were then applied from randomly sampled Gaussian distributions with the voxelwise standard deviation; the absolute resulting signal was taken. Across all voxels, signal noise applied had a mean and standard deviation ($\mu \pm \sigma$) of 5.65±3.21 and 5.51±3.14 for synthetic patients 1 and 2, respectively. The VFA data were recreated using the same signal model and R_{10} maps with FAs of 5°, 10°, 15°, 20°, 25°, and 30°.

Synthetic signal intensity-time data were exported in DICOM file format and the original DICOM DCE-MRI data were replaced by the synthetic data. The patient identifiers were overwritten to avoid confusion with the original RIDER data from which the synthetic data was derived. The parameter maps (K^{trans}, k_{ep}, v_p) were then changed to create the second visit data through the same process with identical AIF. No guarantee was offered that the second visit data were identical to the first, and some substantial differences were deliberately introduced. The differences between visits helped assess accuracy while also penalising methods that over-constrained visits to have repeatable values.

3.2.3 Tumour Segmentation

Segmentation of brain tumours for each visit in the clinical and synthetic data was performed on the last time series of DCE-MRI scans, by comparing to the

anatomical post-contrast T_1 and FLAIR images to delineate the enhancing tumour region. The regions of interest (ROIs) were not released to the challengers. The segmentations were carried out using the ITK-SNAP software (Yushkevich *et al.*, 2016) by an experienced neuroimaging researcher (Reyna Patel) under the supervision of a senior neuroradiologist (Leland S. Hu). These mask data were output in NIfTI format to be overlaid on the submitted K^{trans} NIfTI matrices.

3.2.4 Leaderboard Evaluation

The entry submissions were evaluated using the OSIPI scoring metrics as defined in Table 3.1.

Segmentation Overlay

The segmentation masks were overlaid in Python using the Nibabel library (Brett *et al.*, 2020) onto the ground-truth data and submitted K^{trans} maps for all datasets. The extracted arrays from all submissions were visualised within Python to ensure correct alignment with the segmentation masks. It was found that the NIfTI files submitted had varying alignment quality due to the nature of the analysis techniques in stripping array data from DICOM files. Any submissions with alignment issues were transformed without interpolation using NumPy 90°rotations or axis reflection to ensure full overlap with the correct ROIs within the tumour segmentation masks (Harris *et al.*, 2020). Mean K^{trans} values were calculated by considering the average values – including negative and zero values but excluding NaN values - within the tumour mask ROIs (TM-ROIs) and used within the scoring metrics (Table 3.1).

OSIPI Scoring

The entries were planned to be scored over three main scoring metrics (Table 3.1): accuracy, repeatability, and reproducibility. The three metrics were multiplied to produce a single final score, which implies that a method needs to score well against all three criteria in order to score well overall. Methods should return

values in a repeatable way to allow tracking of any changes which occur and give accurate values for this. Reproducible methods are of pivotal importance to allow dependable use across centres or collaborators. These scoring metrics (Table 3.1) are defined similarly to conventional definitions of accuracy and repeatability. It is worthwhile to mention that we opted for a novel definition of accuracy in this challenge as an alternative to the conventional definition (1-bias), to overcome the limitations such as negative scores. These new metric definitions enable the separate metrics to be combined into one overall score that is equally influenced by the three criteria.

Preliminary Evaluation of the Challenge

An independent team of two scientists, David A. Hormuth II and Julie C. DiCarlo, were invited to perform a test run for the whole challenge process. One of these scientists (David A. Hormuth II) used their in-house K^{trans} quantification software on the DCE-MRI scans of all subjects and visits in the synthetic and clinical cohorts, according to the challenge guidelines. They provided the K^{trans} maps along with SOP of the analysis approach. The second scientist (Julie C. DiCarlo) from the same institution followed the SOP to reproduce the results. The final results by the two scientists were used to test our scoring metrics and revise the challenge guidelines where necessary. As these results may be biased, they are not reported.

3.2.5 Statistical Analysis

To take full advantage of the data submitted for the challenge, submissions were evaluated by several voxelwise approaches; complementing the $Score_{accuracy}$ and $Score_{reproduce}$. This provides vital information, as similar TM-ROI mean K^{trans} values may stem from vastly different distributions (Matejka & Fitzmaurice, 2017).

For accuracy, a voxelwise Bland-Altman analysis was applied to assess the differences between each entry and ground truth K^{trans} values. The mean difference and standard deviation were calculated for all the synthetic visits

Table 3.1: Summary of OSIPI scoring metrics.

Global OSIPI Scoring Metric	
$OSIPI_{gold} = 100 \times (Score_{accuracy} \cdot Score_{repeat} \cdot Score_{reproduce})$ $OSIPI_{silver} = 100 \times (Score_{accuracy} \cdot Score_{repeat})$	
Component Scoring Metrics	
$Score_{accuracy} = \exp\left(-\frac{1}{4}\sum_{n=1}^{4}\frac{\sigma(\bar{K}_i^{trans}, \bar{K}_{exact}^{trans})}{\mu(\bar{K}_i^{trans}, \bar{K}_{exact}^{trans})}\right)$	**Accuracy Score:** A measure for the accuracy of the K^{trans} values of submissions ($\bar{K}_i^{trans}$) by comparison with exact values ($\bar{K}_{exact}^{trans}$) in the synthetic data. The bar here denotes a mean K^{trans} value over the tumour mask in each scan, with averaging over the four synthetic data sets. (σ and μ represent sample standard deviation and mean, respectively, between the two scans.)
$Score_{repeat} = \exp\left(-\frac{1}{8}\sum_{n=1}^{8}\frac{\sigma(\bar{K}_{i,v_1}^{trans}, \bar{K}_{i,v_2}^{trans})}{\mu(\bar{K}_{i,v_1}^{trans}, \bar{K}_{i,v_2}^{trans})}\right)$	**Repeatability Score:** A measure for the repeatability of the K^{trans} values of submissions. This score compares the submitted K^{trans} values for the test ($\bar{K}_{i,v1}^{trans}$) and retest ($\bar{K}_{i,v2}^{trans}$). The bar here denotes a mean K^{trans} value over the tumour mask in each scan, with averaging over the eight clinical patient data sets. (σ and μ represent sample standard deviation and mean, respectively, between the two scans.)
$Score_{reproduce} = \exp\left(-\frac{1}{20}\sum_{n=1}^{20}\frac{\sigma(\bar{K}_i^{trans}, \bar{K}_{i,neutral}^{trans})}{\mu(\bar{K}_i^{trans}, \bar{K}_{i,neutral}^{trans})}\right)$	**Reproducibility Score:** A measure that quantifies to what extent the submitted K^{trans} values are independently reproducible. The metric compares the submitted K^{trans} values ($\bar{K}_i^{trans}$), calculated for the two visits of each of the ten cases (i.e., two synthetic and eight patient data), against the same values reproduced independently ($\bar{K}_{i,neutral}^{trans}$) by experienced neutral evaluators, based on the SOPs provided. The bar here denotes a mean K^{trans} value over the tumour mask in each scan, with averaging over the twenty visit data sets. (σ and μ represent sample standard deviation and mean, respectively, between the two scans.)

combined. Additionally, the proportional change in K^{trans} values within the TM-ROI between the two visits (K^{trans}_{v1} and K^{trans}_{v2}) in synthetic patients was computed.

$$dK^{trans}_{prop} = \frac{K^{trans}_{v1} - K^{trans}_{v2}}{K^{trans}_{v1}} \tag{3.1}$$

For reproducibility, a voxelwise analysis was applied to deduce the differences between the submission and neutral teams' K^{trans}. The mean difference, standard deviation details were calculated for all patient visits combined.

To allow a more detailed conclusion about the repeatability outcomes from the submissions, two metrics were extracted: 1) TM-ROI mean K^{trans} difference between clinical patient visits calculated for each of the 8 clinical patients; 2) repeatability coefficient for TM-ROI mean K^{trans} ($\%RC = 2.77 \times wCV$; where wCV denotes the within-subject coefficient of variation, defined by the root mean square of μ/σ) (Shukla-Dave *et al.*, 2019) to measure repeatability between clinical visits within the same submission.

3.3 Results

Ten submissions, identified by a team name, were received from May 15, 2021 through December 30, 2021. SOPs of four submissions could not be reproduced due to runtime errors or extensively long computational time. Therefore, the reproducibility score was calculated only for six submissions and $OSIPI_{gold}$ was only reported for them. For comparing the methods in terms of accuracy and repeatability, $OSIPI_{silver}$ was used for the remaining submissions, ranked below those with $OSIPI_{gold}$ score.

3.3.1 Overview of Challenge Entries

An overall summary of the procedures used for each submission is provided in Table 3.2, which includes preprocessing methods (brain masking, denoising, co-registration), PK model, AIF selection method, and the DCE-MRI image quantification tool applied. There was a wide variety of AIF selection methods ranging from manual to fully- automatic, but most teams opted to apply the

extended Tofts PK model.

3.3.2 OSIPI Scores

The OSIPI scores, as defined in Table 3.1, for each received entry are shown in Table 3.3 (see Table A.4 for component score confidence intervals). The highest overall score was obtained by the DCE-NET submission with $OSIPI_{gold} = 78\%$, followed by Maydm and PerfLab with 73% and 61%, respectively. $OSIPI_{gold}$ scores ranged from $12 - 78\%$ across submissions with a 59% median score. The $Score_{accuracy}$, $Score_{repeat}$ and $Score_{reproduce}$ ranged from $0.54 - 0.92$, $0.64 - 0.86$ and $0.65 - 1.00$, with median values of 0.69, 0.81 and 0.95, respectively.

3.3.3 Further Evaluation

A summary of the mean K^{trans} values extracted from the TM-ROIs for the clinical patient datasets are included in Table A.1 (Appendix A), with the values for the synthetic patient sets reported alongside the ground-truth in Table A.2 (Appendix A). Figure 3.1 shows the distributions of K^{trans} values within the TM-ROIs for all patients across the submissions. Different methods lead to vastly different mean values and distributions.

An example of the clinical data received from each submission can be seen in Figure 3.2. This details the K^{trans} (min^{-1}) map for both visits in the same slice. The estimations within the TM-ROI and the rest of the brain are highly variable between submissions. A similar plot is detailed in Figure 3.3, where the K^{trans} (min^{-1}) maps are displayed for synthetic patient 2, indicating the variability among different tools. Small visual differences that were picked up by each software between the two visits can be observed.

Accuracy

Figure 3.4 summarises the voxelwise differences between ground-truth and submitted K^{trans} values for the TM-ROI in the synthetic patients. This figure illustrates the wide range of accuracy of the tools in the synthetic data. Some entries (DCE-NET, Madym, ROCKETSHIP and PerfLab) have largely

Table 3.2: An overview of methods used by the challenge entries. Submissions shown using preferred team name with the institution below.

Submissions	Brain mask background removal	Denoising	DCE image co-registration	DCE quantification tool/ Language	AIF selection method	PK model
DCE-NET (Amsterdam UMC)	Yes	No	No	In-house Python (Ottens *et al.*, 2022; Gurney-Champion & Ottens, 2022)	Population- based (Rata *et al.*, 2016)	Extended Tofts
Madym (QBI Manchester)	Yes	No	No	Open-source C++ toolkit DCE-MRI Madym, with its integrated python wrappers (Berks *et al.*, 2021; Berks, 2022)	Population- based (Parker *et al.*, 2006)	Extended Tofts
PerfLab (ISI Brno)	Yes	No	No	PerfLab (Jiřík, 2022)	Automatic (Jiřík *et al.*, 2019, 2022)	Extended Tofts
MRI-QAMPER (MSKCC)	Yes	No	No	MRI-QAMPER MATLAB toolkit	Semi-Automatic	Extended Tofts
FireVoxel (Cornell/NYU)	Yes	Yes, only for VFA dataset	No	FireVoxel/C++ (Mikheev & Rusinek, 2022)	Semi-automatic (sagittal vein)	Extended Tofts
ROCKETSHIP (Barrow)	Yes	Yes	No	ROCKETSHIP (open source)/ MATLAB39	Manual	Extended Tofts
ImageJ/ MRIcron (ADISL)	Yes	Yes	No	ImageJ (Schneider *et al.*, 2012) DCE Module (Open-Source)/ Java	Automatic (Singh *et al.*, 2009)	Extended Tofts
OHSU (OHSU)	Yes	No	No	In-house Python	Semi-automatic	Extended Tofts
UW QBI Lab (Washington)	Yes	No	No	In-house MATLAB	Population-based[a]	Extended Tofts
ALICE (ICL)	Yes	No	Yes	In-house MATLAB	Manual	Tofts, Extended Tofts, Shutter-Speed, No-exchange model

[a]Population AIF extracted from challenge data via manual AIF selection in each challenge patient

Table 3.3: A summary of OSIPI scores for all entries.

Submission	Rank	$Score_{accuracy}$	$Score_{repeat}$	$Score_{reproduce}$	$OSIPI_{silver}(\%)$	$OSIPI_{gold}(\%)$
DCE-NET	1	0.92	0.85	1	78	78
Madym	2	0.85	0.85	1	73	73
PerfLab	3	0.78	0.8	0.98	62	61
MRI-QAMPER	4	0.72	0.86	0.93	62	57
FireVoxel	5	0.57	0.78	0.65	45	29
ROCKETSHIP	6	0.59	0.64	0.74	37	28
ImageJ/ MRIcron	7	0.85	0.68	N/R	58	N/R
OHSU	8	0.67	0.79	N/R	53	N/R
UW QIBlab	9	0.61	0.81	N/R	50	N/R
ALICE	10	0.54	0.86	N/R	46	N/R

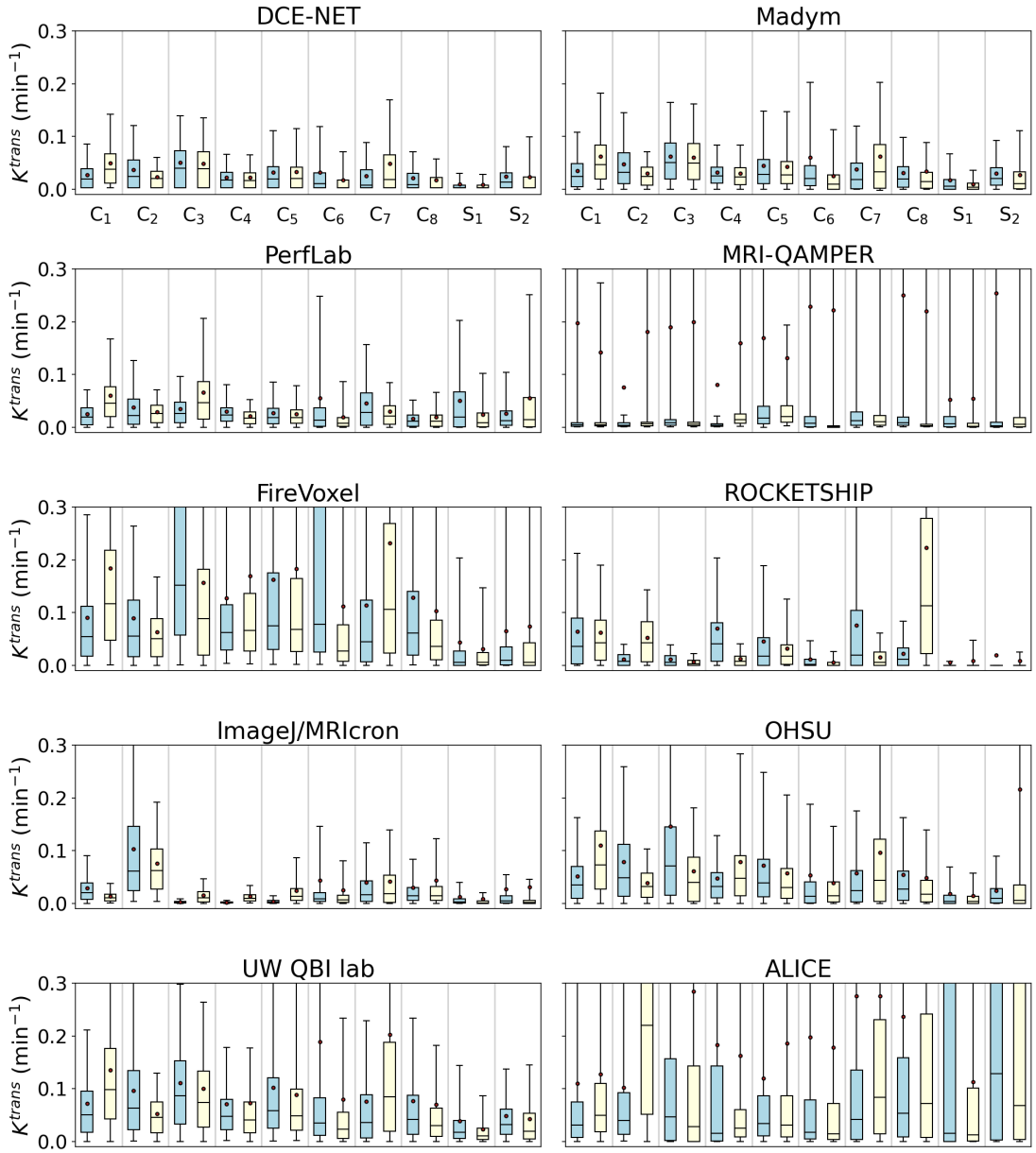

Figure 3.1: Boxplots showing the distribution of K^{trans} values within the tumour mask for each patient visit in each challenge submission. The filled dots denote the mean K^{trans} values over the tumour region with the boxed region and central line showing the interquartile range and median, respectively. Whiskers show the 5th to 95th percentile values. The panels are arranged by the submission team, with grey lines separating each of the clinical patients, and visit 1 (blue) and visit 2 (yellow) both shown. The K^{trans} axis is limited to 0.3 min^{-1} for clarity of comparison.

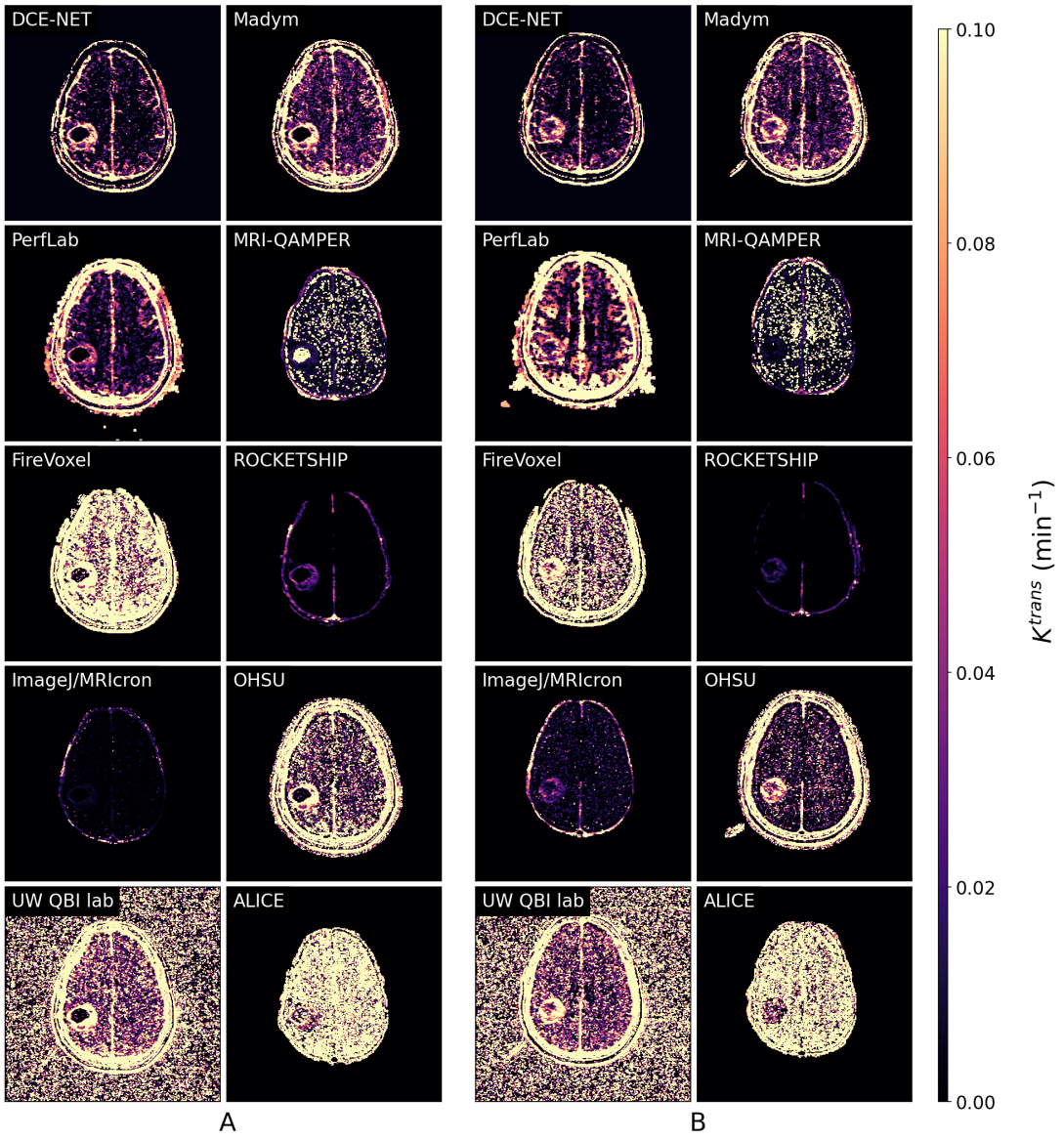

Figure 3.2: K^{trans} values for clinical patient 3 over all submissions. Set A and B correspond to visit 1 and 2, respectively. The maximum K^{trans} value is restricted to 0.1 min^{-1} for comparison. NaN values are set to 0 in this figure for visualization purposes.

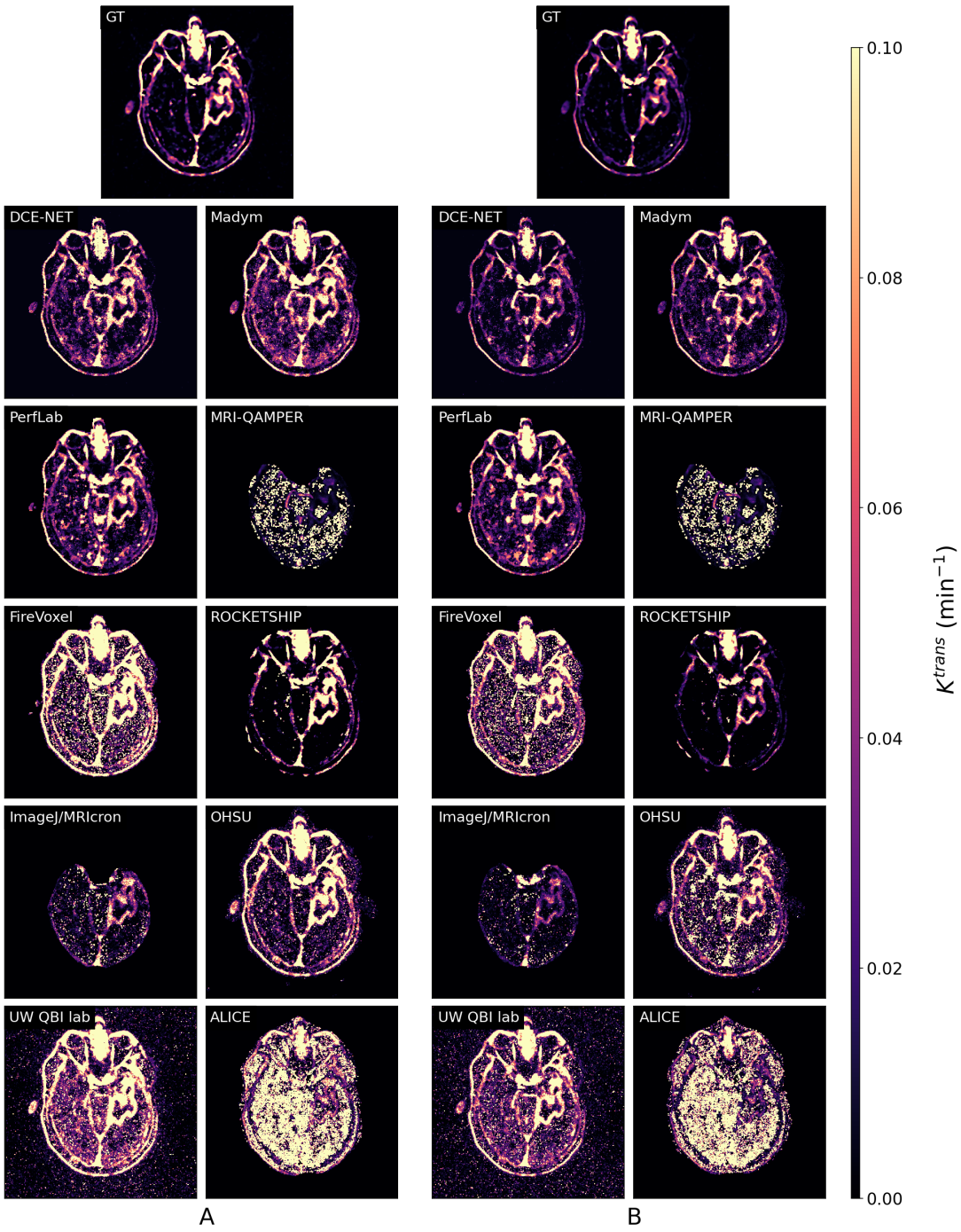

Figure 3.3: K^{trans} maps for synthetic patient 2 for both visit 1 (A) and 2 (B) over all submission teams and the ground truth (GT). The maximum K^{trans} value is restricted to 0.1 min^{-1} for comparison. NaN values are set to 0 in this figure for visualization purposes.

Table 3.4: A summary of the proportional change in mean TM-ROI K^{trans} values (dK^{trans}_{prop}) in synthetic patient 2 for all submissions and the ground truth (GT). To compare the dK^{trans}_{prop} values from each submission, the absolute difference from the GT dK^{trans}_{prop} was calculated.

	Synthetic P_2 (dK^{trans}_{prop})	Absolute difference from GT
GT	0.363	-
MADYM	0.338	0.025
ImageJ/MRIcron	0.331	0.032
UW QBI lab	0.331	0.032
DCE-NET	0.307	0.056
FireVoxel	0.437	0.074
OHSU	0.218	0.145
ROCKETSHIP	0.517	0.154
ALICE	0.089	0.274
MRI-QAMPER	-0.006	0.369
PerfLab	-0.09	0.453

symmetric difference distributions while others (MRI-QAMPER, OHSU, FireVoxel and UW QBI lab) show a tendency to under or overestimate the ground-truth values. Bland-Altman plots of both the TM-ROI mean (Figure 3.5) and voxelwise K^{trans} values (Figure 3.6) reflect trends seen in $Score_{accuracy}$ and Figure 3.4. It should be noted that as Figure 3.4 illustrates, while the $Score_{accuracy}$ from MRI-QAMPER and PerfLab are comparable, the voxelwise K^{trans} values are more variable in MRI-QAMPER, suggesting that the voxels with high values in MRI-QAMPER are averaged out in the calculation of $Score_{accuracy}$ (see Figure 3.7).

A summary of dK^{trans}_{prop} for synthetic patient 2 can be found in Table 3.4. This also includes the absolute difference from the dK^{trans}_{prop} of the synthetic ground truth, ranging from 0.025 to 0.453.

Repeatability

Figure 3.8 shows the distribution of relative changes in K^{trans} values between patient visits, which correspond well with the repeatability score (Table 3.3). A higher mean relative difference of K^{trans} values between the clinical visits corresponds to a lower repeatability score. Within Figure 3.1, the summary of clinical data analysis provides a patient-wise insight on the overall submission

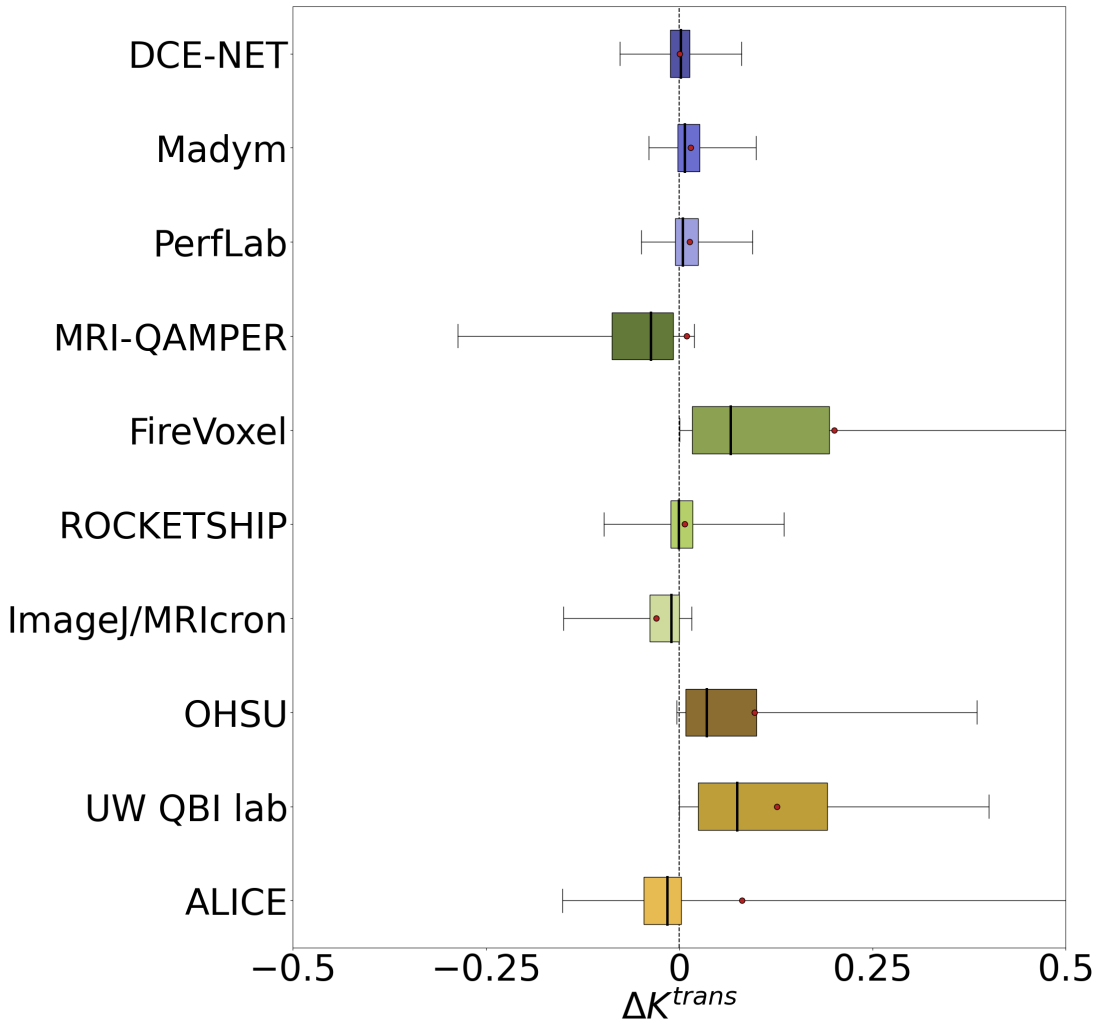

Figure 3.4: Boxplot of voxel-wise differences in K^{trans} values within tumour ROIs between the ground truth and entry values. Filled point shows the value bias (mean difference between entry and ground truth values), with the box and central line showing the IQR and median of the distribution. Whiskers show the 5th to 95th percentile values.

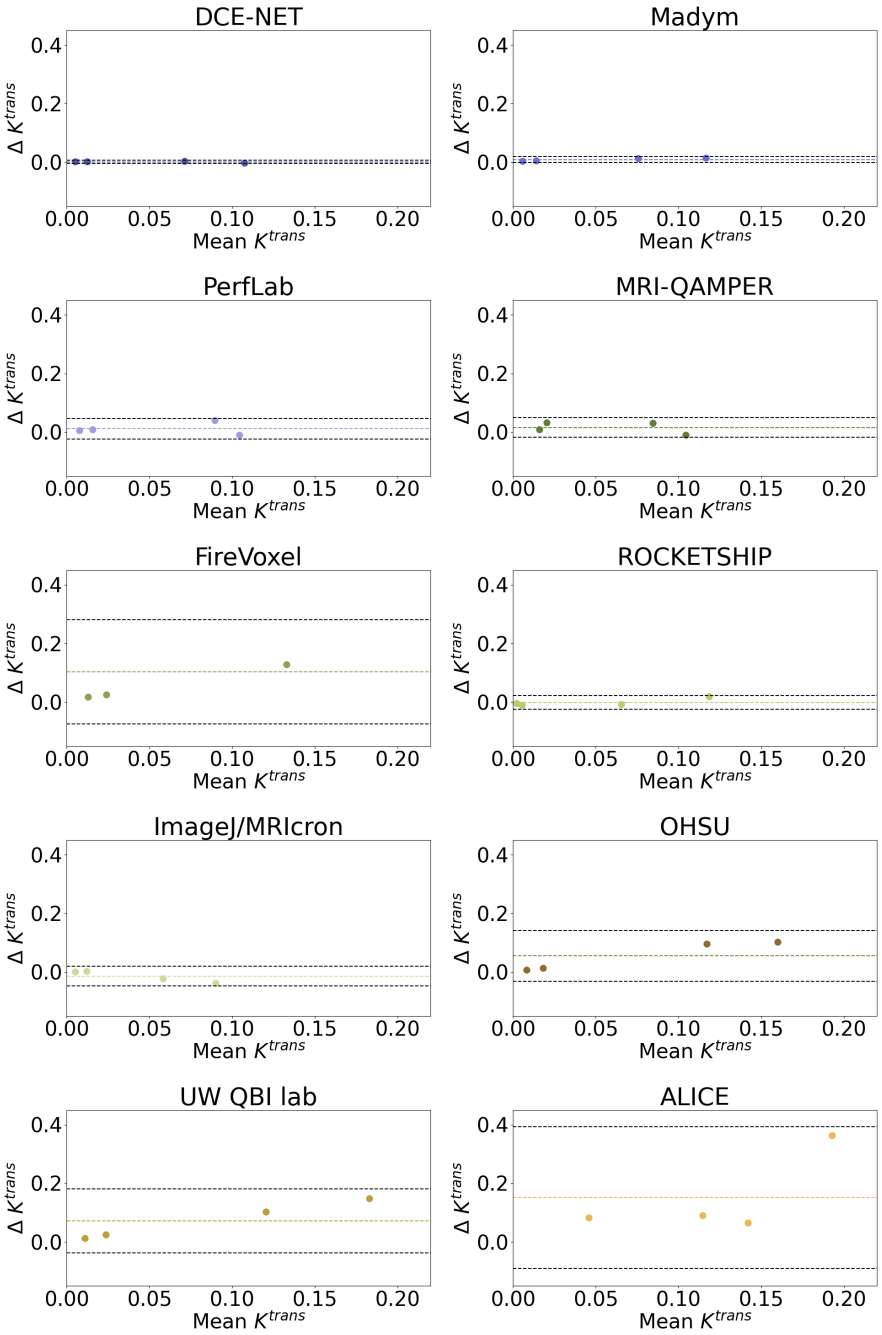

Figure 3.5: Bland-Altman plots showing comparison between the submissions and the ground truth for all synthetic patient visits for mean tumour ROI K^{trans} values. Any zero values within the tumour ROI were included in the analysis. Coloured dashed line and black dashed lines in each panel show the mean difference (bias) and upper/lower limits of agreement (bias$\pm 1.96\sigma$), respectively.

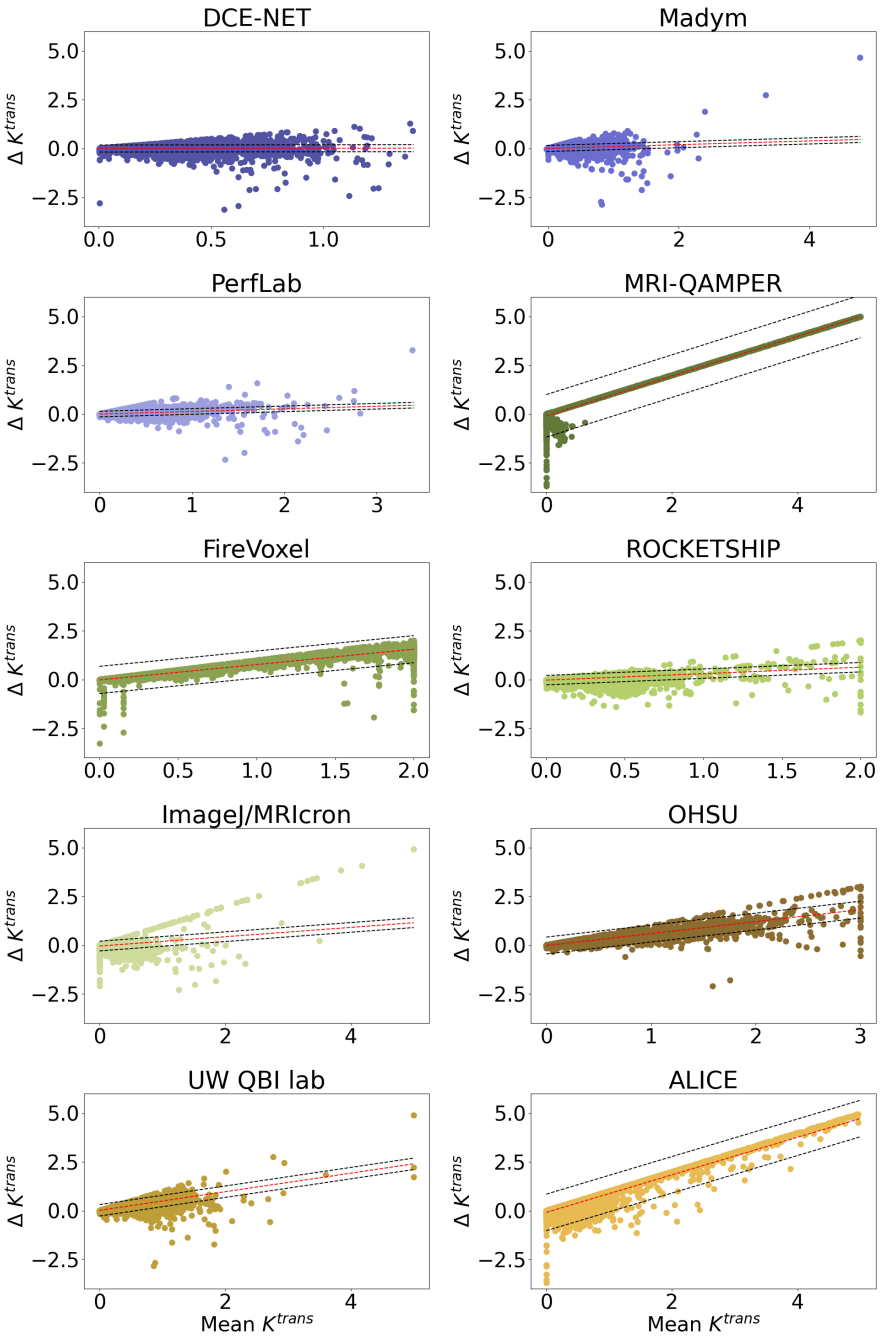

Figure 3.6: Bland-Altman plots showing the comparison between the submissions and the ground truth for all synthetic patient visits for voxelwise K^{trans} values. Any zero values within the tumour ROI were included in the analysis. Red dashed lines show a general linear fit bias with black dashed lines giving the upper/lower limits of agreement (bias$\pm 1.96\sigma$).

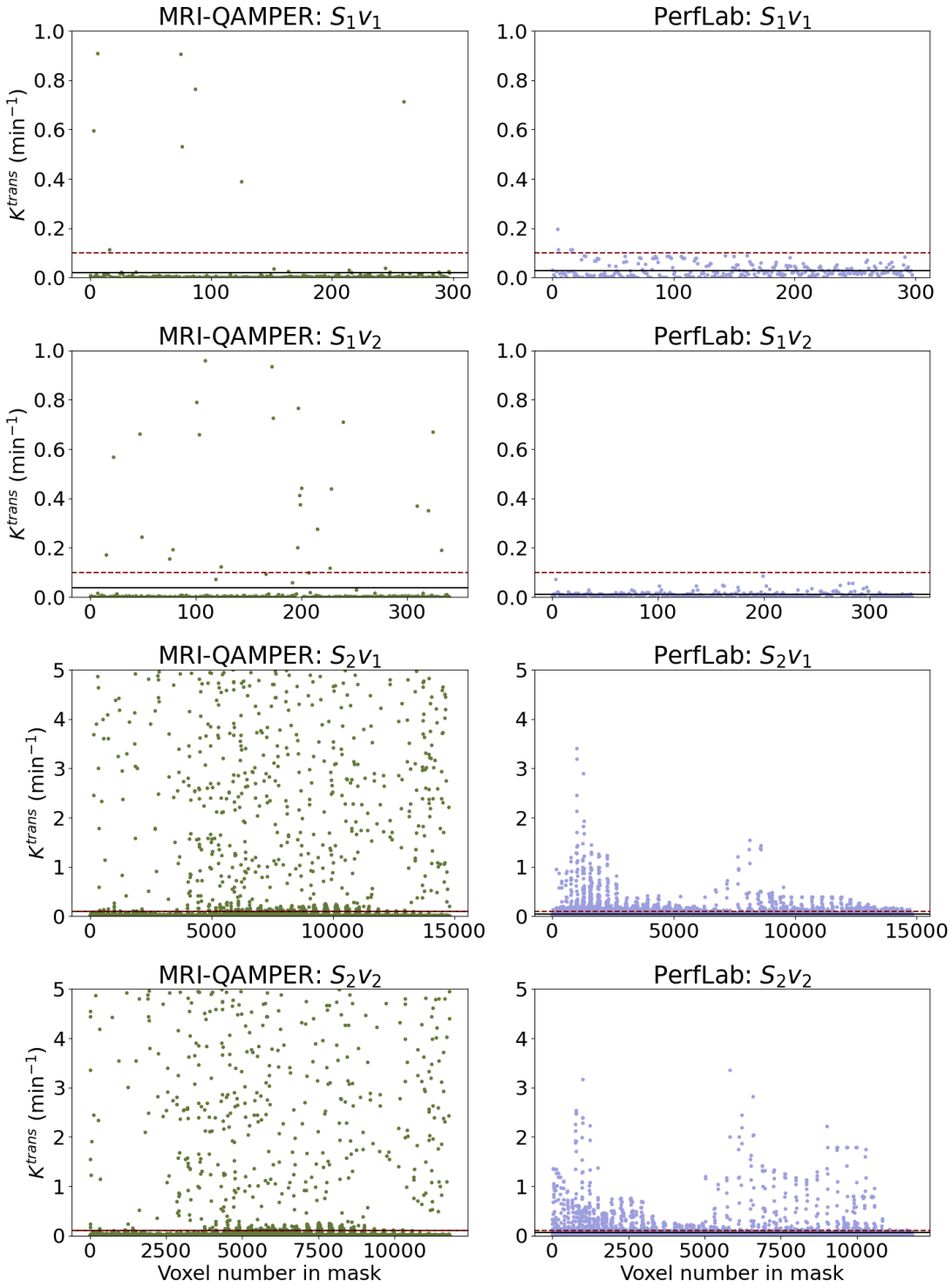

Figure 3.7: Voxelwise K^{trans} values for the synthetic patients across the MRI-QAMPER and PerfLab submissions. In each panel, the black line shows visit tumour mask mean and the red dashed line at 0.1 shows Figure 3.3 cut off value. Different K^{trans} ranges are reported for synthetic patient 1 (top 4 panels) and 2 (bottom 4 panels) to match the range of K^{trans} values present in each patient across both submissions.

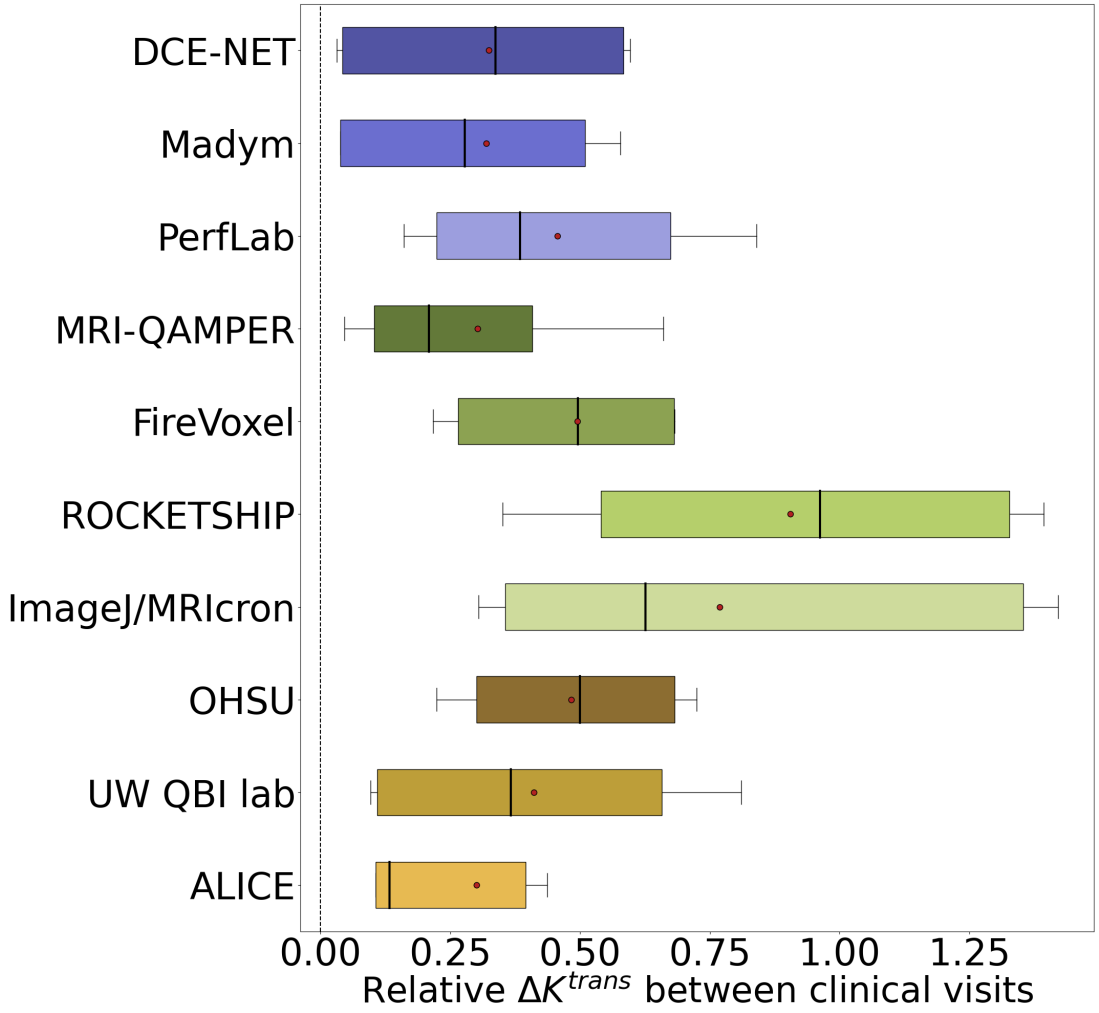

Figure 3.8: Boxplots showing the distribution of absolute relative change in K^{trans} (min^{-1}) values within tumour ROIs
between visit 1 and visit 2 for each submission team. The filled dots denote the mean K^{trans} values over the tumour region with the boxed region and central line showing the interquartile range and median, respectively. Whiskers show the 5th to 95th percentile values.

distributions (Figure 3.8). A Bland-Altman analysis of test-retest TM-ROI mean K^{trans} values (Figure 3.9) reports test-retest variability within submissions.

Table 3.5 shows the %RC values with a range of 0.56-1.45% in the clinical patients comparing the mean TM-ROI K^{trans} between the test and retest visits. The OSIPI repeatability score (Table 3.3) showed strong negative correlation with %RC values between clinical visits, with a Pearson correlation coefficient of -0.986 ($p<0.001$), thereby supporting the validity of the defined repeatability metric (Table 3.1) across these submissions.

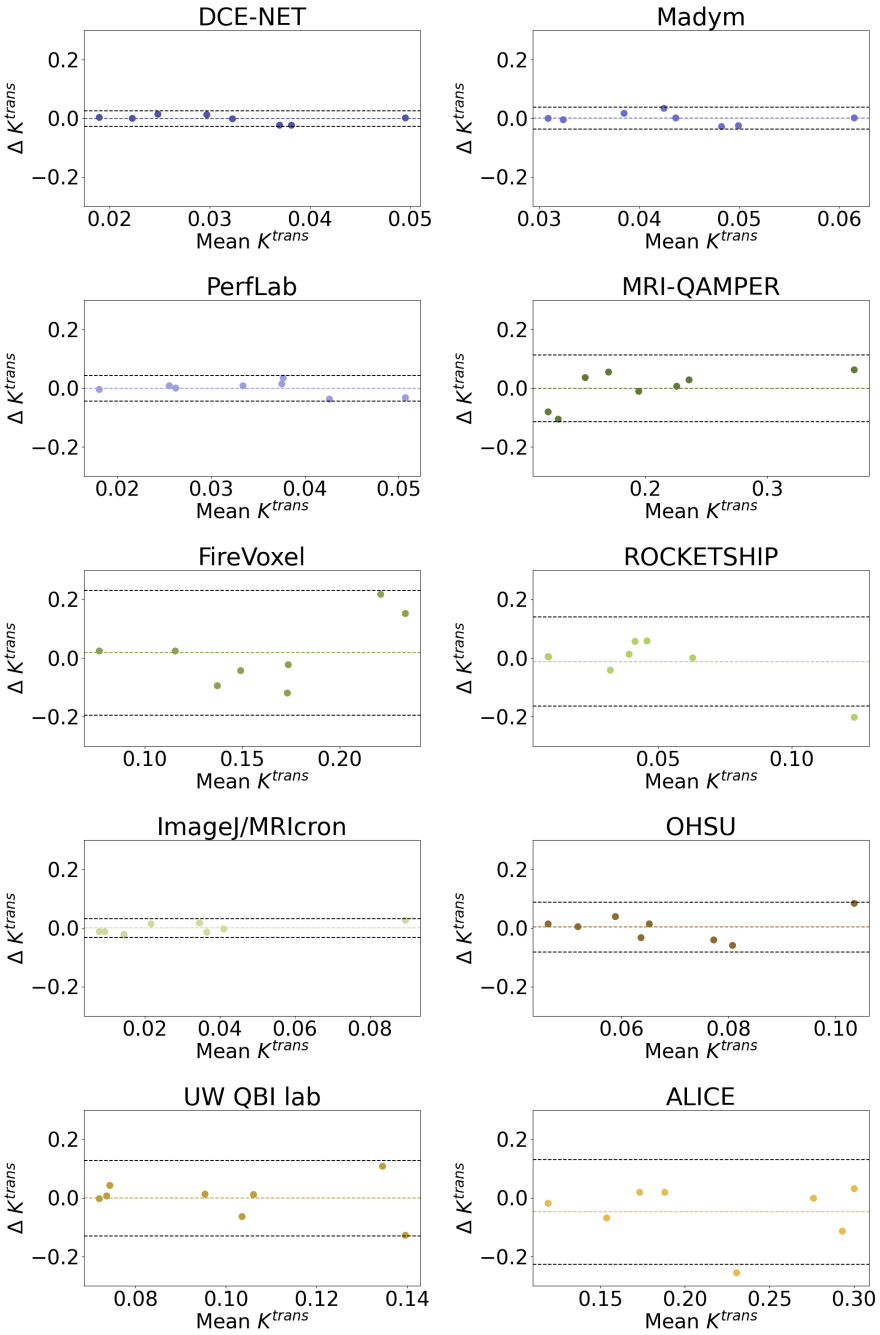

Figure 3.9: Bland-Altman plots showing the comparison between the submissions visit 1 and visit 2 mean tumour ROI K^{trans} values across all clinical patient cases. Any zero values within the tumour ROI were included in the analysis. Coloured dashed line and black dashed lines in each panel show the mean difference (bias) and upper/lower limits of agreement (bias$\pm 1.96\sigma$), respectively.

Table 3.5: %RC values for the clinical test-retest visits, applied to tumour ROI mean values, for each of the submissions.

Submission	%RC	± 95% CI
MRI-QAMPER	0.56	0.26
DCE-NET	0.57	0.24
Madym	0.59	0.27
ALICE	0.62	0.32
UW QBI lab	0.73	0.32
OHSU	0.75	0.23
PerfLab	0.76	0.29
FireVoxel	0.79	0.27
ImageJ/MRIcron	1.29	0.5
ROCKETSHIP	1.45	0.51

Table 3.6: Comparison of original submissions and reproduced K^{trans} values. Columns 1 and 2 shown the mean K^{trans} values across every voxel within masked regions for the original and reproduced results. Columns 3-5 give summary statistics of the differences between the submission and reproduced values calculated within each voxel.

Submission	Mean K^{trans} (min^{-1})		Voxelwise differences K^{trans} (min^{-1})		
	Original	Reproduced	Mean	SD	± 95% CI
Madym	5.4E-2	5.4E-2	-3.9E-13	1.9E-9	9.4E-12
ROCKETSHIP	5.2E-2	5.3E-2	-1.8E-4	1.0E-1	5.2E-4
DCE-NET	4.2E-2	4.2E-2	-1.9E-4	3.2E-3	1.6E-5
PerfLab	4.5E-2	4.4E-2	3.6E-4	2.5E-2	1.2E-4
MRI-QAMPER	1.7E-1	1.7E-1	-5.1E-4	4.7E-1	2.3E-3
FireVoxel	1.6E-1	8.7E-2	6.9E-2	3.1E-1	1.6E-3

Reproducibility

To complement the overall reproducibility scores, shown in Table 3.3, the TM-ROI voxelwise differences between the original and reproduced entry are summarised in Table 3.6 in order of magnitude. The values differ by several orders of magnitude. The ranking in these data largely follows the $Score_{reproduce}$, but for some submissions their rank is slightly improved or worsened. A Bland-Altman plot comparing the TM-ROI mean K^{trans} in each patient visit between the reproduced and original entry (Figure 3.10) follows similar trends to the voxelwise analysis (Table 3.6).

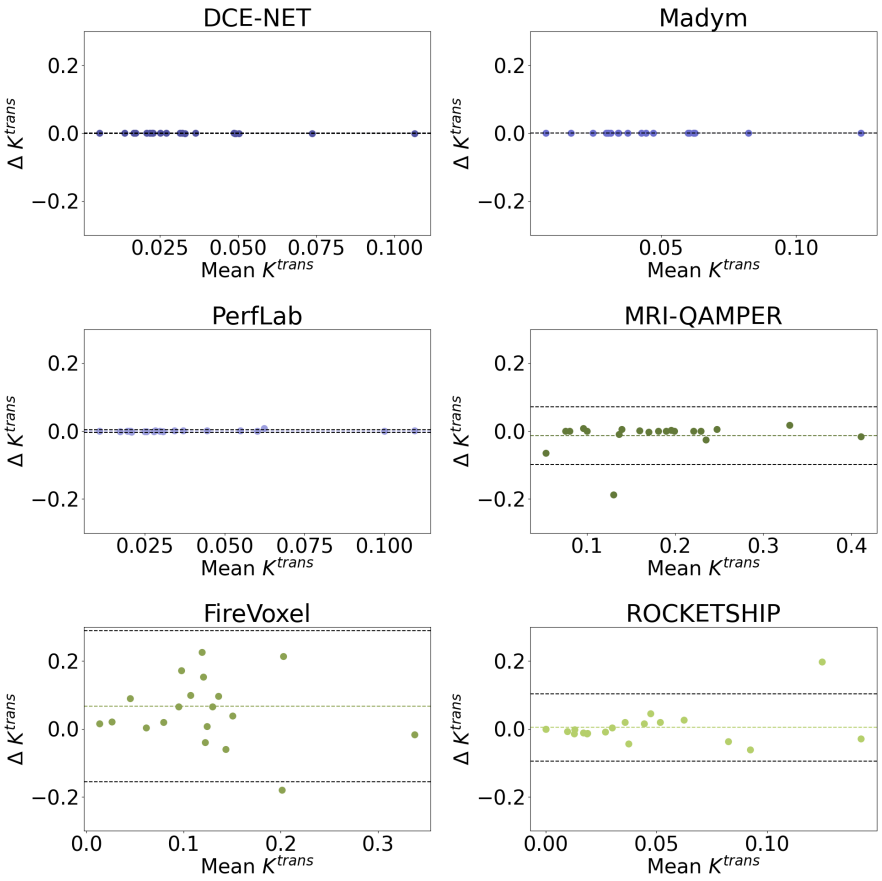

Figure 3.10: Bland-Altman plots showing the comparison between the submissions and the neutral evaluators for all visits mean tumour ROI K^{trans} values. This is displayed only for submissions that were reproduced. Any zero values within the tumour ROI were included in the analysis. Coloured dashed line and black dashed lines in each panel show the mean difference (bias) and upper/lower limits of agreement (bias$\pm 1.96\sigma$), respectively.

3.4 Discussion

In this work, we systematically evaluated variability in quantification of K^{trans} obtained by different analysis pipelines using a standardised benchmark. The submissions were assessed in terms of a scoring model that measured accuracy, repeatability and the ability to reproduce results independently.

3.4.1 Accuracy

Submissions employing population-based AIF (DCE-NET, Madym) scored highly for accuracy. This is interesting as population-based AIFs do not account for between-subject differences and are typically seen as a means of trading off accuracy against precision (Parker *et al.*, 2006; Port *et al.*, 2001). The results indicate that this trade-off is favourable even in terms of accuracy - possibly indicating that AIF measurement biases in this application area are larger than typical between-subject differences. On the other hand, the results may have been biased by the synthetic data generation, for which the maximal value of the widely-used population-based AIF (Parker *et al.*, 2006), though not its functional form, was applied to scale the selected AIF. The other highest scoring methods for accuracy after these submissions were from ImageJ/MRIcron and PerfLab which employed fully automatic AIF methods. Due to the data-driven approach, these software tools should be more robust compared to the methods that utilise population-based AIF in synthetic data that has been developed with a different AIF.

The synthetic data were purposefully different between the 'test' and 'retest' visits to enable detection of any method which attempted to enforce the repeatability between visits if distinct differences in values were present. It should be noted that the synthetic data were not evaluated in the repeatability scoring. This was potentially observed in the ALICE submission, which had a high $Score_{reproduce}$ but a low $Score_{accuracy}$. The ability to detect these changes in K^{trans} was further investigated using dK^{trans}_{prop} to compare to the synthetic data. Based on absolute difference from dK^{trans}_{prop} in the synthetic data, Madym showed the lowest deviation from the known change. This metric aimed to

provide an overview of accuracy less biased to systematic offsets in K^{trans} values which may have resulted from differences in concentration-signal conversion parameters and transit time handling.

Synthetic patient 1 was based on a RIDER patient with no obvious enhancing tumour region, though this information was not revealed to participants; therefore, the ROI was placed within the normal-appearing white matter (NAWM). K^{trans} values in NAWM have been shown to be small but non-zero in several studies at high field strengths (Larsson *et al.*, 2009; Taheri *et al.*, 2011; Cramer *et al.*, 2014), with a distinguishable difference in values also reported between some patient groups with healthy controls. As K^{trans} is expected to be minimal in NAWM, this selection of the synthetic data was meant to evaluate how the analysis tools perform in the absence of blood-brain barrier disruption or the regime of low K^{trans}. This choice could have biased the score against methods that are not optimised to return K^{trans} in NAWM as some of the tools are developed and applied to the enhancing tumour regions. Additionally, methods that cover a wide range of K^{trans} values through estimating more free parameters may overfit a scenario where *a priori* knowledge exists. However, this evaluation was considered necessary for absolute quantification of K^{trans} and standardisation of values across different studies. Closer quantification of low K^{trans} values was observed in synthetic patient 1 in ImageJ/MRIcron, Madym, DCE-NET and PerfLab compared to other submissions (Table A.2); these packages also achieved closer quantification of the tumour region in synthetic patient 2.

3.4.2 Repeatability

In all submissions, %RC values for clinical patients test-retest visits were below the 21.3% threshold (QIBA MR Biomarker Committee, 2020), suggested currently by QIBA as an estimate of true change in assessment of glioblastoma for studies with no repeatability measures. Therefore, with any of these packages, in follow-up studies on treatment response assessment in glioblastoma, any measured changes of K^{trans} that exceed the reported %RC

values (Table 3.5), can be attributed to treatment response with 95% confidence. High repeatability of K^{trans} is essential for longitudinal monitoring of tumour's response to treatment or its progression (Shukla-Dave *et al.*, 2019), as the reconstruction of the same conditions reliably could help show detection of changes should the values deviate. Although high repeatability should not come at the expense of sensitivity to actual changes, this balancing act is crucial for longitudinal studies.

Some or all of the mean TM-ROI K^{trans} values in the MRI-QAMPER, ALICE, and FireVoxel submissions are outside the IQR of the TM-ROI distribution, indicating heavy influence by voxels with outlier values. This may be due to the choice of masks for analysis by these submissions, as some methods are solely designed to return values within tumour regions. If outlier thresholding was applied, these entries may have performed better on this metric as the central IQR distributions appear much more consistent between visits. In spite of this, the repeatability and %RC scores for MRI-QAMPER and ALICE were not overly affected, although this may not hold if the methods were applied to different datasets with more prevalent outliers.

Lower repeatability scores in FireVoxel, ROCKETSHIP, and ImageJ/MRIcron tools may have been associated with their denoising routines within the methodology. Application of this preprocessing step may have influenced differences that have been reported between the visits, potentially applying different amounts of denoising and affecting the resulting parameter retrieval

3.4.3 Reproducibility

The evaluators were able to reproduce six entries with no or limited interactions[1] with the participants. For the remaining entries, interactions consisted of concerns and issues of software malfunction or image processing time. In three out of four teams that remained unreproduced, there were some manual steps involved, namely fully manual (UW QBI lab and ALICE) or

[1]Limited interactions were defined as a small number of interactions on simple issues e.g. resolution of installation problems.

semi-automatic (OHSU) AIF selection.

Specifically, the evaluators encountered software malfunction for ImageJ/MRIcron and OHSU, and incomplete software in UW QBI lab. Finally, due to fitting several PK modelling strategies, long computational time (an estimated ≈100 days) was an issue for the ALICE submission as the timeline available to the evaluators was insufficient to reproduce the results.

Issues highlighted suggest the need for clear guidelines about the level of detail in the SOPs that is required to allow the straightforward replication of the methods for widespread use. SOPs may contain video tutorials walking through each step, clarifying where the installation or runtime errors may stem from. In addition, it would be helpful to combine all software modules into a single executable file so that future users would download the entire software package at once. Besides, it is necessary to minimise any manual decisions that can vary across different operators for higher reproducibility. For example, the software packages requiring manual interaction within the AIF present lower reproducibility in this study, with ROCKETSHIP and FireVoxel receiving the lowest reproducibility scores. In addition, the submissions with the highest reproducibility scores (Madym, DCE-NET and PerfLab) used population- based or fully-automatic AIF selection processes, thereby eliminating most user-specific interactions.

In general, software packages should ideally be developed with community distribution in mind using best practice guidelines (Wilson *et al.*, 2017, 2014; van Houdt *et al.*, 2022), concerning usage guidance, documentation and issue logging. To ensure this, testing with users equipped with a range of expertise and operating systems as well as between institutes is essential. If a software package only runs under a certain system requirements, it should be clearly mentioned in the SOP so that the users can address this prior to installation. Then, the software may be utilized regardless of the users depth of their experience in quantitative DCE-MRI.

3.4.4 Implication on Future Challenge Design

The design of the metric based on the mean K^{trans} values could be biased towards methods with outlier handling, suggesting a score using the median could be more representative of the majority of values produced via these methods. The analysis was rerun using the median K^{trans} values but this did little to change the ranking order; it caused the greatest improvement in accuracy scores from ALICE and FireVoxel.

A second issue encountered was the encoding of NaN values. After preliminary analysis of all submissions, it was discovered that FireVoxel presented values of 1×10^{60} upon extraction with Nibabel package in Python (Brett *et al.*, 2020). After discussion with the authors of this submission, it became clear these were intended to encode NaN values. Therefore, it was decided that these values were treated as NaN and excluded from further analysis. Prior to this correction FireVoxel produced substantially different scores of 0.40, 0.54 and 0.56, for accuracy, repeatability, and reproducibility, respectively demonstrating the importance of proper attention to NaN handling in the challenge design. Inclusion of a specific section within the SOP outlining the NaN handling processes is recommended to avoid this in future, to exclude unphysical or missing values from score calculation. Particularly, masked zero values would artificially lower the TM-ROI mean and may improve $OSIPI_{gold}$ and %RC values.

Our synthetic data was produced using a single-voxel approach. While this approach provides value in terms of benchmarking, giving an equal comparison from known parameters to score the entries, for future work, the use of an interacting voxel simulation would be beneficial (Hanson *et al.*, 2018). This approach would only have been of concern if the entries were hitting perfect accuracy levels suggesting a bias created by the production method. Also, values used for concentration-signal conversions may be beneficial to provide to avoid systematic effects in K^{trans} estimation. Additionally, r_2^* has been shown to vary,(Blockley *et al.*, 2008) but was assumed constant for synthetic data production due to short TE. An interesting extension of the modelling for

future challenges would reduce such assumptions.

For future challenges, scoring should ideally include a reproducibility score for all submitted entries. An alternative setup for reproducibility could be more efficient; perhaps requiring submission of an independent reproduction along with the entry method. However, this may be preventative in the number of received submissions. A checklist for inclusion in the SOP could be of value including a detailed summary of pipeline components, total run time and any licensing requirements.

While our proposed scoring metric for assessment of accuracy measures the bias of the submitted methods in estimating K^{trans} reliably, as noted in the results section, the voxelwise variability in K^{trans} quantification may be averaged out when calculating the TM-ROI mean. Future challenges may account for the voxel-by-voxel differences (instead of TM-ROI mean) between the submitted and ground-truth K^{trans} values.

3.4.5 Study Limitations

The scope of the challenge is limited to K^{trans} and does not necessarily present a full report on the state of the tissue under study. In the case of the extended Tofts model, parameters K^{trans}, v^e and v^p ($k_{ep} = K^{trans}/v^e$) are required to report the full state of the tissue by including the relative volumes of the compartments within the model. The participants were asked to submit K^{trans} values with no requirement to fit any specific PK model, although the extended Tofts model was almost universally applied in the submissions. In future challenges, definition of a PK model to use and requiring submission of all parameter maps (K^{trans}, v_p, k_{ep}) would allow for an improved analysis particularly focusing on co-variance of model parameters. Additionally, requesting AIF details would be recommended to compare the effect of AIF amplitude on variances in parameter estimation between the different approaches. This has been shown to be a factor for K^{trans} estimation (Cheng, 2008) and would inform discussion on the influence of AIF type on parameter accuracy.

The results of this study highlight the variability of pipeline choices in the

submissions received. The presented results and discussions are not able to fully untangle the relative impact of each methodology choice on the resulting K^{trans} values. To address this, future challenges should design methodology to investigate specific pipeline choices. For example, a challenge might supply a smaller data set but ask for several pipeline options.

In this challenge, no commercial software was submitted to the challenge, which was unfortunate and may potentially reflect difficulties in the provision of licenses for the evaluators. Even with the disclaimer that there was no expectation of making code packages or software freely available beyond the evaluators, potentially the open science aspect dissuaded interested parties from the outset. Perhaps including this disclaimer in the advertisement could bring in more commercial packages. Additionally, the risk of reputational damage is speculated to be a factor for the lack of participation by commercial packages. The monetary value of any participating software could be compromised by poor performance against equivalent open-source or free software. Interested parties are encouraged to analyse the OSIPI-DCE challenge data with any commercial packages they hold licenses to enable benchmarking between all software types.

In this study, variabilities in quantitating K^{trans} using different tools were reported. While the submitted tools were ranked using $OSIPI_{gold}$ and $OSIPI_{silver}$, the OSIPI-DCE challenge did not aim to find the "best" tool for analysis of DCE-MRI, rather to provide a platform for comparing the methods. The submitted tools may not have been specifically designed and validated for quantification of K^{trans} in brain gliomas, or they may not have been tailored to the design of the challenge or the specific scoring metrics used. Nevertheless, the proposed $OSIPI_{gold}$ score remains beneficial as a benchmark. Other research groups working on DCE-MRI analysis tools are encouraged to apply their methods to our provided dataset and evaluate their results using our scoring metrics.

These results raise concerns in relation to the accuracy and reproducibility of some software packages, particularly in terms of the validity of their historic application. There is a risk that erroneously fitted values may be present

historically within the field. These may be referred to as literature values, or used to inform choices in future research studies. External reproducibility of modelling remains the gold standard for verification of prior methodologies, and usage of pipelines with proven accuracy, repeatability and reproducibility metrics is essential.

3.5 Conclusion

In conclusion, the OSIPI-DCE challenge highlighted the variability in K^{trans} quantification between submissions and how the choice of methods in analysis pipelines affect K^{trans} estimations. Further developments and consensus are needed within the community to standardise pipeline selection in different clinical settings to estimate K^{trans} at a standard biomarker level. Some aspects that can be improved were identified as: greater detail in description of analysis methodology to enable dissemination of approaches beyond the immediate developers, outlier handling, and the level of manual interactions as in the AIF and brain tissue mask selection. Benchmarking efforts, such as the presented challenge, aid translation of K^{trans} from research to clinical application (O'Connor *et al.*, 2017). Moreover, as the field moves towards increasingly complex PK and signal modelling, and application of deep learning to replace model- based approaches, benchmarking the tools that produce reliable K^{trans} estimations can provide a base for comparison of other advanced markers. To this end, the challenge data and assessment methodology will persist, providing an ongoing benchmarking tool for software development and pipeline selection.

Chapter 4

Review of Spatiotemporal Analysis in Perfusion MRI

As illustrated in the preceding chapter (Chapter 3), analysis of perfusion MRI via typical single voxel methods requires further benchmarking efforts. Further work will help to unravel the causes behind differences observed between different software packages. An avenue of growing interest is the application of more complex models that better represent the underlying physiology of the system being investigated. The application of compartmental modelling principles in a spatiotemporal setting is considered in this thesis. Such modelling efforts have the potential to provide a more accurate characterisation of parameters such as blood flow velocity or perfusion. Additionally, improvement of fitting accuracy from treatment as spatially interacting spaces could lead to enhanced reproducibility, as methods with higher accuracy may provide more reproducible results.

In perfusion MRI, image voxels form a spatially organised network of systems, all exchanging indicator with their immediate neighbours. Yet the current paradigm for perfusion MRI analysis treats all voxels or regions-of-interest as isolated systems supplied by a single global source. This simplification not only leads to long-recognised systematic errors but also fails to leverage the embedded spatial structure within the data.

Since the early 2000s, a variety of models and implementations have been proposed to analyse systems with between-voxel interactions. In general, this leads to large and connected numerical inverse problems that are intractable with conventional computational methods. With recent advances in machine learning, however, these approaches are becoming practically feasible, opening up the way

for a paradigm shift in the approach to perfusion MRI.

This chapter[1] seeks to review the work in spatiotemporal modelling of perfusion MRI using a coherent, harmonised nomenclature and notation, with clear physical definitions and assumptions. The aim is to introduce clarity in the state-of-the-art of this promising new approach to perfusion MRI, and help to identify gaps of knowledge and priorities for future research.

4.1 Introduction

Perfusion MRI includes the subfields T_1-weighted DCE-MRI (Khalifa *et al.*, 2014; Petralia *et al.*, 2020), T_2^*-weighted DSC-MRI (Shiroishi *et al.*, 2015; Boxerman *et al.*, 2020), and ASL (Alsop *et al.*, 2015; van Osch *et al.*, 2018; Hernandez-Garcia *et al.*, 2019). All three methods use an indicator or tracer[2] which modifies the MRI signal in proportion to its concentration - either MR indicators (DCE-MRI or DSC-MRI) or magnetically labelled water (ASL). Rapid dynamic MRI is then used to track the spatiotemporal variations in signal induced by the indicator. After deriving indicator concentration from the measured signal changes, these methods then apply PK models to obtain maps or ROI based measurements of perfusion parameters. This review discusses advanced PK modelling and therefore applies to DCE-MRI, DSC-MRI and ASL alike.

The conventional approach to PK modelling in perfusion MRI describes the concentration in each voxel or ROI independently by a 1D (temporal) PK model. Conceptually this builds on the fundamental assumption that each voxel or ROI acts as an isolated system with a single, global inlet of indicator (Tofts & Kermode, 1991; Tofts *et al.*, 1999; Sourbron & Buckley, 2012, 2013; Petralia *et al.*, 2020). The concentration in the inlet is typically assumed to be known and referred to as the AIF. This assumption effectively separates the problem of modelling a single large 4D dataset into a large number of small and

[1] The work in this chapter has been accepted for publication as (Shalom *et al.*, 2024a), this is defined in 'Joint Publications' on Page ii. The text has been edited to conform with UK spelling conventions, and replaces a singular reference to 'this paper' with 'this review'.

[2] Used in ASL.

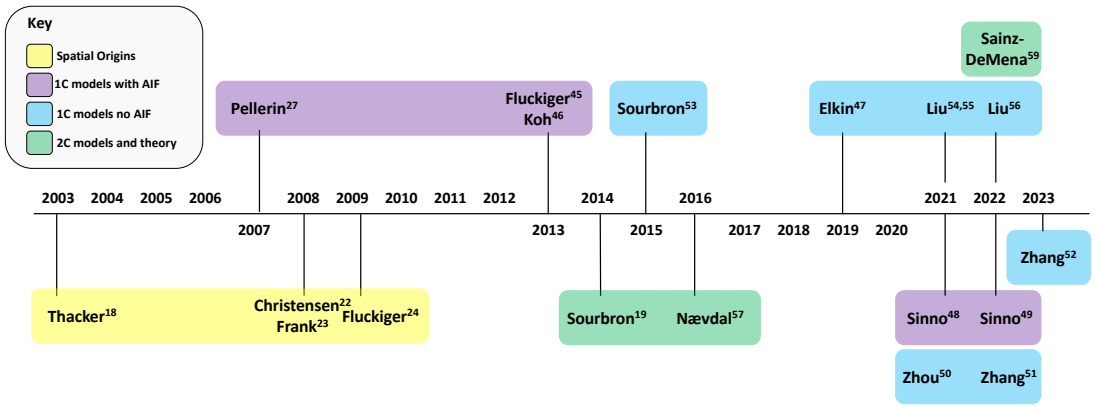

Figure 4.1: Timeline of contributions within the literature landscape leading towards developing spatiotemporal TKs. The studies listed are grouped by the theme of the work or model applied using distinct colours as indicated by the key.

independent 1D problems. This makes the analysis highly scalable, parallelisable, and computationally efficient. On the other hand, the assumption is obviously invalid and it has been known for over 20 years that this leads to significant systematic errors (Calamante *et al.*, 2000; Buckley, 2002; Calamante, 2013; Willats & Calamante, 2013; Hanson *et al.*, 2018; Petralia *et al.*, 2020).

In principle, the problem can be resolved by dropping the isolated-systems assumption and modelling all voxels in the imaged volume as connected systems that all exchange directly with their neighbors (Henkelman, 1990; Thacker *et al.*, 2003; Sourbron, 2014). Unfortunately, this approach presents significant computational challenges that have so far proven insurmountable. Yet with the increase in computational power and the advance of machine-learning (Raissi *et al.*, 2019, 2020), solutions are becoming practically feasible. New approaches to spatiotemporal modelling of DCE-MRI are increasingly proposed, but comparing methods and models between papers presents a significant challenge due to differences in physical concepts, terminology and notations.

The aim of this review is to summarise all relevant developments on spatiotemporal PK modelling of perfusion MRI data in a common framework. This will establish a firm foundation for future developments, facilitate identification of knowledge gaps, and lower the barrier for entry in the field for new researchers.

4.2 History and scope

The discussion about the foundations of perfusion MRI is as old as the field itself. In 1990, Henkelman (1990) argued that the very definition of perfusion as inflow per unit volume is not physically justifiable in an imaging setting because fluid flow scales with area rather than volume. While undeniably true, these objections were largely ignored until the early 2000's when Thacker *et al.* (2003) presented a technique for DCE-MRI that introduced the concept of cerebral blood flow orientation in terms of spatial gradients in mean transit time. Similar ideas were proposed by Christensen *et al.* (2008) in DSC-MRI using the gradient of arterial delay times to derive information about the directionality and orientation of perfusion. Within ASL, an experimental approach for measuring perfusion orientation was proposed which involved labelling planes in different orientations (Frank *et al.*, 2008). From a different angle, the idea of spatial coherence between neighbouring voxels has been exploited to estimate AIFs from tissue-level data by a joint fitting of multiple voxels (Fluckiger *et al.*, 2009; Jiřík *et al.*, 2019, 2022).

While these ideas go some way to demonstrate the potential of using spatial information in the analysis, they are limited by the lack of a clear underlying theoretical framework that can be used to build models of spatiotemporal indicator propagation. Since the early 2010's increasing numbers of papers have proposed concepts borrowed from continuum mechanics, computational fluid dynamics (CFD) or porous media theory to build spatiotemporal generalisations of classic 1D PK models. The earliest proposal dates back to Pellerin *et al.* (2007) using a spatial model of intervoxel diffusion but retaining the concept of a global AIF to model indicator delivery to the voxel through the vasculature. A first step towards a more general formulation, albeit conceptually confused, can be found in a self-published report from 2013 (Pautot, 2013).

In 2014 these ideas were placed on a more rigorous footing including also multi-compartmental systems (Sourbron, 2014). One of the consequences is the natural emergence of a formal definition of perfusion (F, in units of mL/min/mL)

as the divergence of the arterial flow ($\mathbf{f}^a$, in units of mL/min/cm^2):

$$F = -\nabla \cdot \mathbf{f}^a \tag{4.1}$$

In words, this states that the perfusion of a piece of tissue is the part of the arterial flow into the tissue that is converted into venous flow out of the tissue. In particular, this does not include the contribution of blood vessels (arterial or venous) that pass through the tissue without feeding its capillaries. This formal definition therefore correctly formalises the true physiological notion of perfusion as 'feeding flow' or 'capillary flow' into a given tissue. Moreover, since divergences scale with volume, this fully resolves Henkelman's original objection (Henkelman, 1990) to the conventional definition of tissue perfusion as inflow per unit volume.

The main barrier to a more widespread adoption of spatiotemporal models is the computational challenges in applying these to the *inverse* problem of deriving perfusion parameters from data. While the number of free parameters *per voxel* is similar to standard 1D models, the voxels can no longer be solved independently. Spatiotemporal models therefore present a single global inverse problem with, for a typical 3D time series, millions of free parameters. And while the problem is linear for the simplest 1-compartmental spatiotemporal models, it is non-linear in the more general setting of multi-compartment models.

The scope of this review is therefore restricted to studies applying spatiotemporal modelling to the inverse problem. This excludes a substantial body of literature using spatiotemporal models as forward models to simulate data. Often these papers aim to generate digital reference objects to investigate and quantify the error caused by neglecting spatial coherence. An example is a study by Barnes *et al.* (2014) investigating the impact of intra-voxel diffusion on the accuracy of conventional DCE-MRI parameters. Another example is a model of the circulation designed to determine the accuracy of conventional perfusion analysis as a function of ROI size (Hanson *et al.*, 2018). Incidentally, the latter provides *in-silico* support for Henkelman's objection, showing that the classic definition of perfusion creates a dependence on voxel size, with increasing bias for smaller voxels. Another application of CFD-type forward models is to

help predict and understand drug delivery to tissues, informed by DCE-MRI data (Steuperaert *et al.*, 2019; Wu *et al.*, 2020, 2021). Often these models are multi-scale, coupling flow in large blood vessels to microvascular flow and interstitial transport. This is part of a wider literature on multi-scale computational modelling of the circulation and biological transport mechanisms (Peyrounette *et al.*, 2018). While these models may be informed by perfusion MRI, they are out of scope for this review unless the models are used to fit the spatial perfusion parameter fields from measured data.

4.3 Model classification

A timeline of the publications in scope for this review is shown in Figure 4.1. Comparison of the model architectures described in these papers reveals 9 nested spatiotemporal models of increasing complexity, illustrated in Figure 4.11 and defined using harmonised notations in Table 4.1. These 9 nested models can be classified as either one-, two- or three-compartment models depending on the number of distinct compartments in each tissue voxel. Within each group they can be further differentiated based on (1) the transport mechanisms described, such as diffusion, convection, or exchange; (2) the compartment types such as interstitial, arterial, or venous space; and (3) whether an external input function is utilised.

The symbols and notations in this review have been modified from the original publications and harmonised as shown in Table 4.1 following the definitions in (Sourbron, 2014). The aim is to reveal the structural differences and similarities between models more clearly. The total tissue concentration, $C(\vec{r},t)$, is a directly measurable quantity and is defined as number of CA molecules per volume of tissue (mmol/mL). If the tissue is built up of multiple compartments, the contribution of a compartment γ to the tissue concentration is given as $C^{\gamma}(\vec{r},t)$, and physically defined as the number of CA molecules in the compartment γ, relative to the volume of the entire tissue (mmol/mL). Examples are the tissue concentration in interstitium $C^{e}(\vec{r},t)$, plasma $C^{p}(\vec{r},t)$, arteries $C^{a}(\vec{r},t)$ or veins $C^{v}(\vec{r},t)$. With this definition, the total tissue

concentration $C(\vec{r},t)$ is always the direct sum of the concentrations in the individual compartments. For instance, if a tissue is modelled as consisting of plasma and interstitial compartments, the total tissue concentration is:

$$C(\vec{r},t) = C^p(\vec{r},t) + C^e(\vec{r},t) \tag{4.2}$$

The volume fractions do not appear in these equations because concentrations are defined relative to the total tissue volume rather than the compartmental volume. While the equations can be recast to an alternative picture involving volume fractions, flows, and perfusion explicitly (see (Sourbron, 2014) for details), this introduces additional free parameters that then have to be constrained by adding new constraints. The total blood flow per unit surface area ($\mathbf{f}$, in units of mL/min/cm^2) is defined from velocity ($\mathbf{u}$, in units of cm/min) and volume fraction (v, in units of mL/mL), by $\mathbf{f} = v\mathbf{u}$. As $\mathbf{f}$ is incompressible the systems are constrained as:

$$\nabla \cdot \mathbf{f} = v\nabla \cdot \mathbf{u} + \mathbf{u} \cdot \nabla v = 0 \tag{4.3}$$

Using tissue concentrations up front simplifies the equations and numerical challenges, and ensures the models are defined using the least number of free parameters. After solving for the models in this picture, any missing markers such as volume fraction, perfusion or blood flow can then be derived as described in (Sourbron, 2014).

The indicator in a compartment γ is transported *between* voxels by velocity fields, $u^\gamma(\vec{r})$, and diffusion fields, $D^\gamma(\vec{r})$. Indicator exchange within a voxel between compartments β and γ is denoted by rate constants $K^{\gamma\beta}(\vec{r})$, describing exchange from β to γ. Some models have an AIF, $c_{in}(t)$, without a positional coordinate $\vec{r}$, that represents a global vascular input.

The following sections provide a more detailed description of the 9 models identified including numerical implementations, and important results. For clarity, the models are detailed in order of increasing model complexity.

4.4 One-compartment models

In one-compartment systems, the compartmental indicator concentration is simply the tissue concentration and is therefore directly accessible through measurement. The implications are that one-compartment models with convection and diffusion can be recast as a first-order linear system of equations. The majority of the work on inverse approaches for spatiotemporal models has focused on one-compartment systems. These effectively describe the voxel as a single compartment with a uniform concentration, and model the exchange of indicator between voxels using diffusion and convection either separately or concurrently. Beyond the precise type of contrast mechanisms, these models differ in the physical compartment that is modelled (intravascular or extravascular), or, equivalently, which tissue spaces are assumed to carry negligible amounts of indicator.

An increasing body of evidence using forward models has demonstrated that ignoring between-voxel interstitial exchange can lead to significant bias on parameters such as K^{trans} (Egeland *et al.*, 2008, 2011; Cho *et al.*, 2009; Egeland *et al.*, 2012; Barnes *et al.*, 2014; Woodall *et al.*, 2018). Initial developments in spatiotemporal analysis of perfusion MRI therefore aimed to eliminate this bias by modelling interstitial convection and diffusion. Additionally, the introduction of interstitial convection enables the accurate representation of tumour regions with significant interstitial fluid pressure gradients that drive detectable advective transport (Boucher *et al.*, 1990; Milosevic *et al.*, 2001; Zachos *et al.*, 2001; Milosevic *et al.*, 2014). Only more recently, attention has turned to spatiotemporal models of vascular transport by convection and diffusion.

4.4.1 Interstitial diffusion

Pellerin *et al.* (2007) introduced a one-compartment model with interstitial diffusion and a global vascular input function (Figure 4.4(a)). The model introduces a new interstitial diffusion parameter, D^e, which acts to transport CA through the interstitium between adjacent voxels. The model does not incorporate vascular transport between voxels, instead retaining the assumption

that CA is delivered to the voxel through a global AIF. Effectively the model, therefore, extends the standard Tofts model(Tofts & Kermode, 1991; Tofts *et al.*, 1999) with between-voxel diffusion in the interstitial space. As in the standard Tofts model it is assumed that the concentration in the plasma space is negligible compared to that in the interstitial space, so the measured concentration is made up of the interstitial concentration only ($C = C^e$).

Pellerin *et al.* (2007) reduce the computational challenge of the inverse problem by considering a 2D system only, and assuming D^e everywhere is a known parameter. All D^e values were fixed to a constant (in the simulations) or spatially dependent but derived from a measured apparent diffusion coefficient (ADC) of water (in data). Optimisation was implemented using a simulated annealing algorithm (Aarts & Korst, 1989), a stochastic optimisation method that improves parameter recovery in systems with local minima (Kirkpatrick *et al.*, 1983).

Experiments include a synthetic data set and 2D slices of mouse DCE-MRI data. The synthetic data modelled 2D circular tumour with a highly perfused rim and necrotic core. K^{trans} was defined to be zero at the core so the only means of CA transport in the core is via interstitial diffusion. The conventional Tofts model produced unphysical $v_e > 1$ in the necrotic core, underestimated K^{trans} in the rim, and overestimated it in the core. These biases disappear after adding the interstitial diffusion terms, allowing a non-vascular transport pathway to the tumour core.

In 2013 (Fluckiger *et al.*, 2013), the same team reduced the computational complexity of the model by assuming the differences in v^e and the diffusion coefficient between adjacent voxels are negligible. These assumptions decouple the equations of individual voxels, allowing for a voxel-by-voxel analysis of the data. This also implies that the interstitial diffusivity of the CA can be fitted as a free parameter.

The result is a drastic reduction in computation time compared to the original model from Pellerin *et al.* (2007): on the same reference object and using the same optimisation method, computation time was reduced from 70hrs to 52s - almost reaching the efficiency of standard Tofts modelling (11s). Unfortunately, the

results also showed large spatial gradients in diffusivity and K^{trans}, indicating that the assumptions do not capture the true behaviour of the system. Additionally, due to the voxel-wise fitting approach the control over global CA conservation is eliminated. This drawback is recognised within the work (Fluckiger *et al.*, 2013), and it is proposed that future iterations of the method should seek to enforce global mass conservation.

Also in 2013, Koh *et al.* (2013) proposed a one-compartment model with interstitial diffusion (Figure 4.2(a)). This simplifies the model proposed by Pellerin *et al.* (2007), by removing CA delivery through a global input followed by extravasation. As such, between-voxel diffusion remains the sole mechanism for indicator transport through the system. The model equations can in principle be solved directly for the diffusion coefficient by dividing the time-derivative of the concentration by its Laplacian. Experiments for this method (Koh *et al.*, 2013) included 14 sets of 3D mice xenograft DCE-MRI data of varying cancer types. In practice, stability in the presence of noisy data was improved by clustering voxels with similar contrast-enhancement patterns, and solving for a single diffusion coefficient in each cluster. While this approach is obviously limited by the strong assumption of diffusion-only transport, it presents an elegant solution for areas such as homogeneous necrotic tumour cores where these assumptions are justified.

4.4.2 Interstitial convection and diffusion

In 2019, Elkin *et al.* (2019) introduced a one-compartment model with interstitial convection and diffusion (Figure 4.3(a)). The model applies interstitial diffusion and convection parameters to distribute CA through the interstitium between adjoining voxels. A global vascular input is not included in the model, effectively assuming that all transport between voxels takes place via the interstitium.

Elkin *et al.* (2019) reduce the scale of the inverse problem by asserting the CA mass density can be written as a function of velocity. For optimisation, an operator splitting method followed by a Gauss-Newton minimisation is applied (Steklova & Haber, 2017). Uniquely, a forward flux is defined as the average

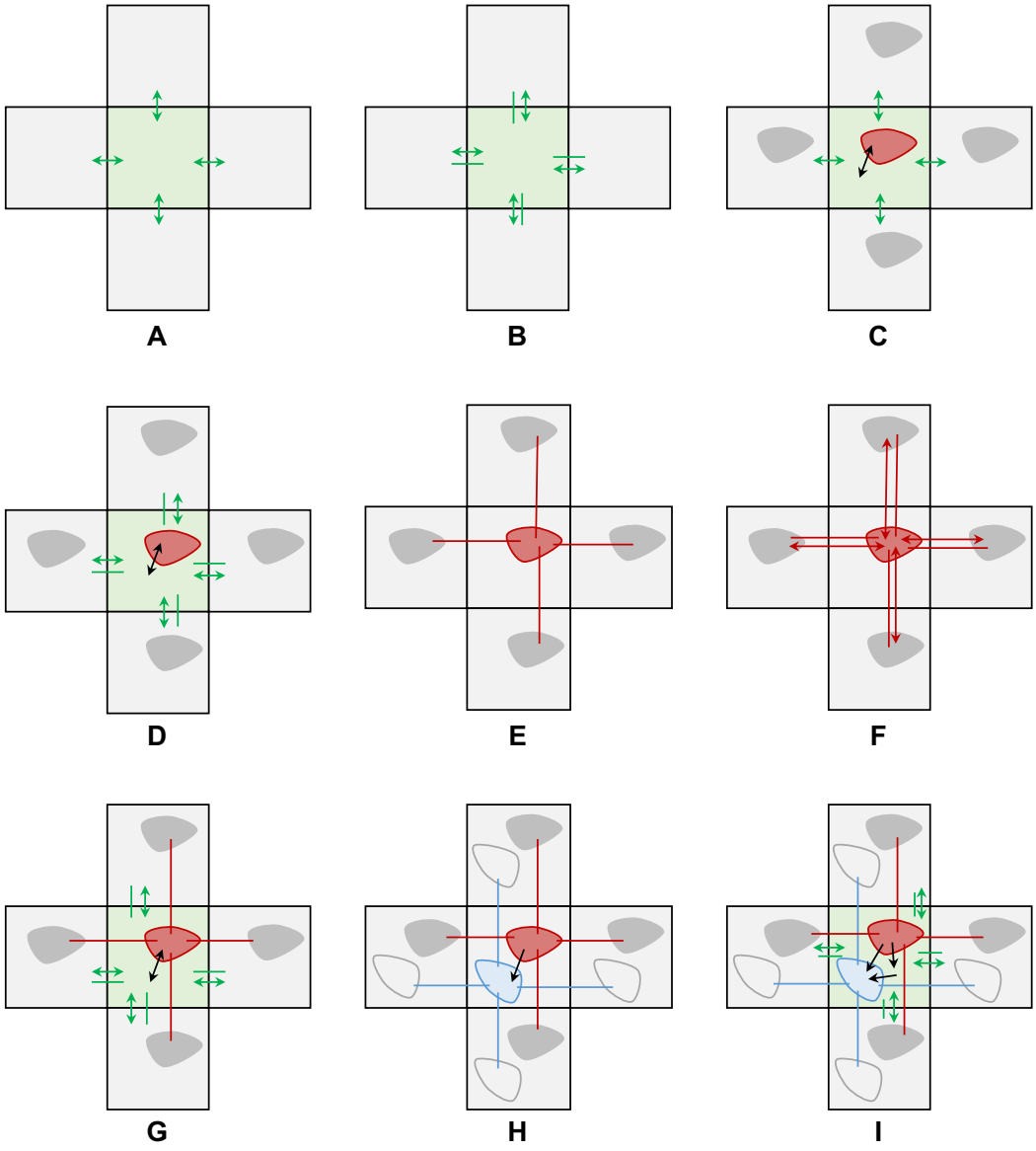

Figure 4.11: Diagrams of the 9 spatiotemporal models proposed in the literature (model equations are given in Table 4.1). Each model is illustrated for a central voxel and 4 neighbors, with interstitial (green) and/or vascular compartments (red for arterial or total blood compartments, and blue for venous). Solid coloured lines and double-ended arrows show between-voxel transport by convection and diffusion, respectively, within a given compartment. Black arrows show within-voxel exchange between different compartments. Shown are (A) A one-compartment system with interstitial diffusion; (B) A one-compartment system with interstitial convection and diffusion; (C) A one-compartment system with interstitial diffusion and a vascular input; (D) A one-compartment system with interstitial convection and diffusion and a vascular input; (E) A one-compartment system with vascular convection; (F) A one-compartment system with vascular convection and diffusion; (G) A two-compartment system with vascular convection, interstitial convection and diffusion with bidirectional exchange; (H) A two-compartment system with vascular convection and mono-directional exchange; (I) A three-compartment system with interstitial convection and diffusion, vascular convection, and directional exchange.

Table 4.1: Summary of general system types for spatiotemporal TKs. Models are named from the transport mechanisms and input type. All models are defined diagrammatically in Figure 4.11, with each specific sub-figure indicated here. The transport equations are shown for each model. Studies from the literature concerning the theory or implementation of specific model types are detailed. Cases that utilise different but equivalent equations from the presented models are tagged with a star (⋆). Where implementations are available the dimension of the method is shown, any models tagged N/A are purely theoretical.

Model	Fig	Dim	Studies	Transport Equations
1C interstitial diffusion	4.2(a)	3D	(Koh *et al.*, 2013)	$\frac{\partial C^e}{\partial t} = D^e \nabla^2 C^e$
1C interstitial convection/ diffusion	4.3(a)	3D	(Elkin *et al.*, 2019)	$\frac{\partial C^e}{\partial t} = D^e \nabla^2 C^e - \nabla \cdot u^e C^e$
1C interstitial diffusion with vascular input	4.4(a)	2D	(Pellerin *et al.*, 2007; Fluckiger *et al.*, 2013)	$\frac{\partial C^e}{\partial t} = K^{trans}\left(c_{in} - \frac{C^e}{v_e}\right) + \nabla \cdot D^e \nabla C^e$
1C interstitial convection/ diffusion with vascular input	4.5(a)	2D	(Sinno *et al.*, 2021, 2022)	$\frac{\partial C^e}{\partial t} = K^{trans}\left(c_{in} - \frac{C^e}{v_e}\right) + \nabla \cdot D^e \nabla C^e - \nabla \cdot u^e C^e$
1C vascular convection	4.6(a)	3D	(Zhou *et al.*, 2021; Zhang *et al.*, 2023)	$\frac{\partial C^p}{\partial t} = -\nabla \cdot u^p C^p$
1C vascular convection/ diffusion	4.7(a)	2D 3D	(Sourbron, 2015; Zhang *et al.*, 2022; Liu *et al.*, 2021a,b, 2022)	$\frac{\partial C^p}{\partial t} = -\nabla \cdot u^p C^p + \nabla \cdot D^p \nabla C^p$
2C interstitial convection/ diffusion with vascular convection and exchange	4.8(a)	N/A	(Sourbron, 2014)	$\frac{\partial C^p}{\partial t} = -\nabla \cdot u^p C^p - K^{ep} C^p + K^{pe} C^e$ $\frac{\partial C^e}{\partial t} = -\nabla \cdot u^e C^e + \nabla \cdot D^e \nabla C^e + K^{ep} C^p - K^{pe} C^e$
2C vascular convection and exchange	4.9(a)	2D	(Naevdal *et al.*, 2016)⋆	$\frac{\partial C^a}{\partial t} = -\nabla \cdot u^a C^a - K^{va} C^a$ $\frac{\partial C^v}{\partial t} = -\nabla \cdot u^v C^v + K^{va} C^a$
3C interstitial convection/ diffusion with vascular convection and exchange	4.10(a)	N/A	(Sourbron, 2014)	$\frac{\partial C^a}{\partial t} = -\nabla \cdot u^a C^a - K^{va} C^a - K^{ea} C^a$ $\frac{\partial C^v}{\partial t} = -\nabla \cdot u^v C^v + K^{va} C^a + K^{ve} C^e$ $\frac{\partial C^e}{\partial t} = -\nabla \cdot u^e C^e + \nabla \cdot D^e \nabla C^e + K^{ea} C^a - K^{ve} C^e$

velocity magnitude over an initial time period and a backward flux for the remaining time interval.

Experiments included 10 sets of 3D head and neck squamous cell carcinoma patient DCE-MRI data. While both the standard Tofts model K^{trans} parameter and proposed forward flux follow similar trends, abrupt changes between neighbouring slices and voxel are present in the K^{trans} maps. Accounting for between voxel transport helped maintain the integrity of the forward flux estimations in the same regions.

Recently, Sinno *et al.* (2021) explored parameter recovery in tumour regions using a one-compartment system with interstitial convection and diffusion with a global vascular input. This model is an extension of the standard Tofts model with additional diffusion and convection terms for CA transport through the interstitium between neighbouring voxels (Figure 4.5(a)). For vascular input, the model applies an AIF-based approach. It, therefore, extends the approach proposed by Pellerin *et al.* (2007) with interstitial convection, or generalises the model in Elkin *et al.* (2019) with a global AIF.

Sinno *et al.* (2021) reduce the complexity of the inverse problem by assuming radial symmetry, fitting 1D ROIs extending from the tumour centre. Within each ROI diffusion is further assumed to be constant. The optimisation approach is largely standard, applying a MATLAB® nonlinear solver.

Experiments detail a set of 2D synthetic radially symmetric tumour models (Sinno *et al.*, 2021), and a set of 2D human cervical carcinoma xenograft DCE-MRI data (Sinno *et al.*, 2022). Their results highlighted a successful differentiation of increased tumour periphery velocities along with considerable diffusivity at the tumour core. Due to assumed symmetry, indicator flow is restricted to along the radial direction only. Within the xenograft study v_e is assumed constant, in contrast to their synthetic study where it was a free parameter. A sensitivity analysis showed no evidence of an impact on transport parameter fits for a fixed v_e between 0.5 and 1. However, for smaller fixed v_e the results became significantly different. As such, the validity of fixing v_e is highly dependent on the influence it exerts on its co-variant parameters.

4.4.3 Vascular convection

Several studies from a group at Cornell have developed a one-compartment model with vascular convection (Zhou *et al.*, 2021; Zhang *et al.*, 2023). The model introduces a spatially variable velocity coefficient, u^p, which acts to transport CA through the vascular space between adjacent voxels (Figure 4.6(a)). Any diffusive transport between voxels is neglected due to the large magnitude of blood velocity.

For the inverse problem, this group (Zhou *et al.*, 2021; Zhang *et al.*, 2023) uses least-squares optimisation on the concentrations to fit for u^p in each voxel. Specifically, an alternating direction method of multipliers with a conjugate gradient algorithm is applied. A regularisation term based on the velocity gradient is employed, acting to enforce smoothness in the recovered velocity field.

Experiments comprise 3D synthetic data sets and 3D clinical data covering varied physiologies such as liver (Zhang *et al.*, 2023) and kidney (Zhou *et al.*, 2021). For synthetic data production, a 1D non-linear network of cylindrical models - solved using Poiseuille's law - are employed to represent the 3D microvascular network. To compare against the ground truth the convection and Navier-Stokes velocities are assumed to be equivalent.

For the synthetic data sets in Zhou *et al.* (2021) the introduced method achieved a smaller u^p error than the Kety's method blood flow when compared with the ground truth u^p values. While these approaches are clearly limited by an assumption of convection-only transport, for intravascular indicators or highly vascularised well-mixed systems this may well be justified.

4.4.4 Vascular convection and diffusion

The Cornell group, applying a similar inverse approach (Zhou *et al.*, 2021; Zhang *et al.*, 2023), developed their method to include diffusive transport (Zhang *et al.*, 2022). Experiments include 3D clinical breast DCE-MRI data, where Zhang *et al.* (2022) reported a more statistically significant distinction between malignant and benign breast tumours in u^p than K^{trans} from the Tofts model.

Sourbron (2015) introduced a one-compartment model with both vascular convection and diffusion (Figure 4.7(a)). The inclusion of vascular diffusion increases the number of free parameters per voxel compared to convection alone, but actually simplifies the numerical problem by allowing for bi-directional exchange at every voxel interface.

With this generalisation, the inverse problem becomes linear and can be solved with standard matrix inversion methods. The unknowns of the discrete inverse system are rate constants at each voxel surface, which represent a combination of the diffusive and convective transport parameters. After solving the linear system for these rate constants, the results can then be converted back to convection and diffusion fields.

Experiments include a 2D synthetic data test case with a population AIF at selected boundary voxels, and the transport equations are solved by forward propagation of the linear system. Results showed that while the concentrations were reconstructed accurately from the data, the fitted parameter maps showed a deviation from the ground truth. These results indicate that the inverse problem in spatiotemporal modelling of DCE-MRI is not in general well-posed and multiple possible solutions exist that are compatible with the data. Strategies to resolving the degeneracy include refining the experimental conditions (e.g. faster injections or sampling), and/or adding regularising constraints to select solutions with particular properties.

Within the DSC community, Liu *et al.* (2021a) also propose a one-compartment model with both vascular convection and diffusion (Figure 4.7(a)). The model introduces spatially variable velocity and diffusion coefficients, u^p and D^p, respectively, which act to transport CA through the vascular space between adjacent voxels.

To reduce the complexity of the estimation Liu *et al.* (2021a) assume u^p is incompressible. In consequence, this effectively constrains the system to have a constant volume fraction (Equation 4.3). For the inverse problem, Liu *et al.* (2021a) use a stochastic gradient descent method to minimise the mean square error between the model and measurement concentration. Regularisation terms based on gradients in diffusion and velocity are employed to enforce smoothness

in recovered transport parameter fields. Their approach returns 3D maps of u^p and D^p.

Experiments consist of 2 synthetic and 43 human stroke lesion DSC-MRI data sets, both in 3D. Their synthetic data sets comprise of (1) purely convective and (2) purely diffusive transport, with various noise levels. The ground truth maps used for u^p are derived using the inverse technique on a brain DSC-MRI data set, while D^p are derived from ADC values.

Their synthetic investigations report a low error recovery of u^p and D^p that is robust to noise level increase. Within the stroke lesion study, Liu *et al.* (2021a) consistently report lower velocity and diffusion values within lesion regions than normal regions. Additionally, their feature maps report a similar or improved interpretation of the stoke lesions when compared to standard perfusion maps.

The same group utilise partially supervised Convolutional Neural Networks (CNNs) fitting the same system type (Figure 4.7(a)) to decrease computational time (Liu *et al.*, 2021b, 2022). These methods (Liu *et al.*, 2021b, 2022) also apply velocity incompressibility and parameter regularisation as in the original study (Liu *et al.*, 2021a). Across 10 of the same ischemic stroke data sets the new CNN-based methodology showed greater distinction between lesion and normal regions than their previous work (Liu *et al.*, 2021a) or standard perfusion metrics.

4.5 Multi-compartmental models

While one-compartment models have some practical utility, it is well-known that most tissues require at least two compartments for an accurate description of their indicator concentrations. For instance, the assumption that intra- and extravascular spaces are well-mixed, is in general not justified. Unfortunately, moving from one-compartment to multi-compartment spatiotemporal models comes with a step change in computational complexity.

In a multi-compartment setting, the concentrations in the individual compartments are hidden and only the total concentration is directly accessible to measurement (Equation 4.2). Hence the multi-compartment spatiotemporal model inherently requires solving a non-linear system with hidden variables, or

a linear system of higher order.

The literature is extremely sparse. Most spatiotemporal equivalents of standard multi-compartment models only exist as theoretical proposals, or still rely on a global temporal input function, which does not model between-voxel transport in the vasculature.

4.5.1 Interstitial diffusion and vascular input

Sainz-DeMena *et al.* (2022) report a two-compartment system with interstitial diffusion and a vascular input. This model applies a diffusion coefficient, D, which acts on the total tissue concentration. The vascular component is a non-negligible plasma space v_p with spatial variation and supplied by a global AIF.

Sainz-DeMena *et al.* (2022) reduce the computational complexity of the inverse problem by considering 2D systems only, and assuming D everywhere is a constant known parameter. Minimisation was implemented using a Trust Region Reflective algorithm, which handles sparse matrices efficiently.

Experiments include the 2D circular synthetic tumour previously proposed (Pellerin *et al.*, 2007; Fluckiger *et al.*, 2013) and a 2D heterogeneous synthetic tumour with various noise levels. The proposed diffusion term enabled the method to consistently outperform the extended Tofts model for parameter accuracy in noise-free scenarios. For systems with low measurement noise, the method showed significantly reduced fitting accuracy, particularly for v_p, compared to the relative stability of the extended Tofts model.

4.5.2 Vascular convection with exchange

Naevdal *et al.* (2016) implemented a two-compartment system defined by Sourbron (2014), modelling blood flow in arterial- and venous compartments and mono-directional transport from arteries to veins by perfusion. Their implementation employs a Darcy flow approach to define the arterial and venous velocities, with the intravoxel exchange from artery to vein mediated by a porous capillary space. Darcy flow is commonly used in porous media to describe pressure-driven fluid flow. (Whitaker, 1986). To relate this model to

biology, porosity and permeability are interpreted in terms of compartmental volume fractions and the transport between compartments.

Concerning the inversion problem, Naevdal *et al.* (2016) decrease the computational complexity by applying *a priori* knowledge of either the permeabilities or porosity values to reduce the number of free parameters. For optimisation, an Ensemble Kalman filtering method is applied, a popular method for parameter estimation in geoscience (Iglesias *et al.*, 2013).

Experiments included two synthetic 2D systems. The proposed method was applied for 2 separate investigations, either using known porosity values, or known permeability values. While the accuracy of the recovered porosity and permeability values are encouraging, the special cases presented apply very specific assumptions that would be inaccessible from a clinical DCE-MRI data set.

4.5.3 Three-compartment models

A theoretical three-compartment system has been proposed(Sourbron, 2014) characterised by separate arterial and venous compartments with convective transport and an interstitial compartment with both convective and diffusive transport. These compartments interact with a mono-directional exchange from artery to vein or interstitium and interstitium to vein (Figure 4.10(a)), following the picture of microvascular exchange involving extravasation at arterial ends of capillaries and reabsorption at venous ends. All the previously presented lower complexity compartment models are special cases of this general description. To the best of the authors' knowledge, there currently exists no implementation of a three-compartment system.

4.6 Discussion

This review has presented 9 nested compartmental approaches that currently exist within the community. Of these 9 models, 7 have existing numerical implementations to recover between-voxel transport coefficients, covering

systems from pure interstitial diffusion to dual vascular convection with exchange. The presented approaches differ in complexity and applied assumptions, but all seek to extract spatial information that is inaccessible to single voxel modelling.

While all approaches build in methods of transport between voxels, either by diffusion and/or convection, some still assume a global vascular input to supply each voxel (Pellerin *et al.*, 2007; Fluckiger *et al.*, 2013; Sinno *et al.*, 2021, 2022). While convenient, this in some sense bypasses the key challenge of modelling transport to a voxel via exchange with neighbouring voxels. Most recent work therefore has focused on removing the assumption of a global source (Koh *et al.*, 2013; Sourbron, 2015; Zhang *et al.*, 2022, 2023). Looking forwards, the further development of methods that do not require a global input is critical to achieve realistic models of indicator propagation across larger distances.

A prevailing problem limiting progression within this topic, is the availability of software implementations from previous studies. To the best of the authors' knowledge, there are no freely available software implementations for any of the presented methods. Consequently, in order to apply or develop any of the previously implemented approaches, researchers are faced with the major challenge of replicating the synthetic data and inversion methodology. Such re-implementation is a significant time investment and acts as a barrier to the future development of otherwise promising methodologies. Moving forwards, increased efforts to publish algorithm and software details via open-source sharing platforms such as GitHub would be invaluable. Not only will open science enable fast external implementation of existing methods but it can also help boost citations and collaboration opportunities (McKiernan *et al.*, 2016). In recent years, there has been increased focus on open science within the perfusion imaging community via the forming of the International Society for Magnetic Resonance in Medicine Open Science Initiative for Perfusion Imaging (ISMRM-OSIPI) an initiative and activity of the ISMRM Perfusion study group (OSIPI, 2023). Contributions from OSIPI and related projects cover challenges, code libraries (van Houdt *et al.*, 2023), standardised data formats (Clement *et al.*, 2022), and recommended lexicon naming conventions for DCE, DSC and

ASL (Dickie *et al.*, 2023; OSIPI Taskforce 4.1, 2023). Application of these and other software development guidelines (Wilson *et al.*, 2014, 2017; Hunter-Zinck *et al.*, 2021) to new contributions within the field will help to accelerate the pace of progression.

Another major hurdle to the development of useful spatiotemporal TKs modelling is the runtime of newly developed methods. Moving from single voxel modelling where each voxel may be fit independently, to a scenario where all voxels must be concurrently fit requires increased computational power. Some of the newly developed techniques apply assumptions within their models to reduce the number of free parameters per voxels, to reduce the computational requirement. While a useful exercise, new developments should focus on a reduction of physically inaccessible assumptions. Presenting an overview of computational runtimes for the implementations discussed in this review is not feasible without replicating the studies: apart from two studies (Pellerin *et al.*, 2007; Fluckiger *et al.*, 2013) reported runtimes are unavailable.

Spatiotemporal TK analysis would benefit from a fully generalised method for parameter reconstruction of any specified compartment model from tissue concentration data. Going forward, research in the field should focus on the development of methods that tackle the multi-compartment inversion problem from small *in-silico* test systems up to 4D *in-vivo* data sets. This development to increasingly complex systems will incur a heavy computational load. For example, a fully spatial model with just two compartments has up to 5 free parameters to fit per voxel - compartmental convection and diffusion coefficients and an inter-compartment exchange term. Such a high number of target parameters alongside a large volume of data appears to lend itself to machine learning approaches, such as the CNN methods proposed (Liu *et al.*, 2021b, 2022). A relatively new branch of the machine learning field is PINNs (Raissi *et al.*, 2019; Sharma *et al.*, 2023), which incorporate the underlying system physics within the loss function to avoid unphysical solutions. The most applicable advance from the PINNs field is a method developed to identify parameters of the Navier-Stokes equations from concentration-time data (Raissi *et al.*, 2020). A promising future direction for this work would be the

adjustment of the PINNs network architecture to handle compartmental structures and TKs equations. Such a network would need to output the compartmental concentrations alongside relevant transport coefficients. To construct the physics-informed aspect of the network, governing equations would be specified (e.g. any system in Table 4.1) and used to form residual equations. Appropriate steps for non-dimensionalisation would be required, alongside suitable activation functions and weighting schemes (Jagtap *et al.*, 2020; Wang *et al.*, 2020). Due to the PINNs layout, modification of the output fields and system dimensions should be relatively straightforward, thereby creating a general inversion framework.

On a fundamental theoretical level, the spatiotemporal field currently lacks some broader understanding of the uniqueness of solutions, and to what extent this is affected by experimental conditions. This has been identified as a problem in several studies (Sourbron, 2015; Sinno *et al.*, 2021), and is pivotal to the future development of inversion methods. Especially in multi-compartmental systems, proof of unique solutions would increase confidence in results where recovered parameters show good agreement with concentration data. The extent to which an AIF is recoverable from the available measurement data, or whether it needs to be separately measured, is of particular interest. Additionally, further investigation into the dependence of uniqueness on experimental design is needed to reliably define solvable systems and conditions. Similar work on system design in standard perfusion quantification demonstrates that solutions degenerate if the indicator is not injected rapidly, or if sampling is too slow or too limited in duration (van Osch *et al.*, 2003; Klawer *et al.*, 2018). It is likely that similar limitations are valid for spatiotemporal models, but no data currently exists to guide experimental design.

4.7 Conclusions

Nine nested model architectures for vascular-interstitial tissues have been identified, although two of those have only been described theoretically. The most complex model currently implemented is a spatiotemporal

two-compartment exchange model. While these developments show promise, there exist unmet needs for model assumptions that apply to real-world problems and for robust computational approaches to the inverse problem.

Chapter 5

Identifiability of Spatiotemporal Tissue Perfusion Models

Over the last decade, there has been building interest in new methods for fitting spatiotemporal TK parameters DCE-MRI data to replace typical single-voxel approaches. The previous chapter (Chapter 4) summarises contributions and advances made within the spatiotemporal field. This overview identifies several gaps within the research landscape, mainly the need for inverse methods focused on two-compartment systems. This chapter presents a proof of concept for fitting the kinetics of 1D spatial systems with one and two compartments, without assuming a measured AIF.

For each of the two models, identifiability is explored theoretically and in-silico for three systems. Concentrations over space and time are simulated by forward propagation. Different levels of noise and temporal undersampling are added to investigate sensitivity to measurement error. Model parameters are fitted using a standard gradient descent algorithm, applied iteratively with a stepwise increasing time window. Model fitting is repeated with different choices for the initial values to probe uniqueness of the solution. Reconstruction accuracy is quantified for each parameter by comparison to the ground truth. This chapter[1] reports positive outcomes in recovering arterial inputs and compartmental parameters.

[1]The work in this chapter has been published (Shalom *et al.*, 2024c), this is defined in 'Joint Publications' with contribution details on Page ii. Figures from the Supporting Information are included in the main thesis text.

5.1 Introduction

Conventional PK modelling of perfusion imaging typically quantifies indicator transport parameters using an isolated single-voxel approach. Although these methods are highly scalable and computationally efficient, they suffer from the fundamental assumption that each voxel acts as an isolated system with a known global inlet concentration - the AIF. This approximation leads to significant model errors which increase with spatial resolution (Buckley, 2002; Calamante, 2013; Willats & Calamante, 2013; Hanson *et al.*, 2018).

In theory, this bias can be removed by the use of spatiotemporal PK models (Sourbron, 2014). Implementations of this approach have mainly focused on one-compartment models with transport by diffusion (Koh *et al.*, 2013), convection (Zhou *et al.*, 2021; Zhang *et al.*, 2022, 2023), or both (Elkin *et al.*, 2019; Sourbron, 2015). Hybrid approaches have also been proposed, coupling a one-compartment spatiotemporal model for interstitial transport with vascular delivery modelled by a single, global AIF (Pellerin *et al.*, 2007; Fluckiger *et al.*, 2013; Sinno *et al.*, 2021, 2022; Sainz-DeMena *et al.*, 2022). Experience with fully spatiotemporal two-compartment systems is extremely limited (Naevdal *et al.*, 2016).

All the above implementations apply additional constraints on the reconstructed model parameters, for instance an assumption that diffusion is constant in space (Pellerin *et al.*, 2007), that the diffusion gradient between adjacent voxels is negligible (Fluckiger *et al.*, 2013), that parameter fields have small spatial gradients (Zhou *et al.*, 2021; Zhang *et al.*, 2022, 2023; Liu *et al.*, 2021a), that transport is only radial in a lesion (Sinno *et al.*, 2021, 2022), that perfusion is modelled by Darcy flow Naevdal *et al.* (2016), or that parameter fields are in a known relationship to each other Naevdal *et al.* (2016). Constraints of this type are included to reduce the computational complexity, but it is not always clear that they are physically justified, creating a risk of new biases. For instance, it has been suggested that Darcy flow (Whitaker, 1986) cannot capture the complete physiology of the capillary bed without further modification (Peyrounette *et al.*, 2018).

Previous studies did not investigate whether such additional constraints, and

the biases they produce, are actually necessary. A simple parameter-counting exercise suggest they might not be: a 3D spatiotemporal model is massively overdetermined, with the data points vastly outnumbering the unknowns. The aim of this chapter, therefore, was to investigate whether unconstrained spatiotemporal PK models are fundamentally identifiable. The study addresses the question in 1D toy models for intravascular indicator where solutions can be generated efficiently with simple optimisation routines.

5.2 Methodology

The question of identifiability is investigated for two nested models of increasing complexity: a one-compartment convective blood flow model, and a two-compartment perfusion model with convective blood transport Sourbron (2014). The perfusion model was selected as it presents challenging conditions for model fitting due to the relative similarity between the two compartments (arterial- and venous).

The whole pipeline from forward modelling to inversion method is contained within the `tkspace` python package (Shalom, 2023). This is available in a GitHub repository[1] containing example scripts used to produce data detailed in this paper and general usage information.

5.2.1 Theory

One-compartment blood flow model

The one compartment blood flow model is defined by the transport equation

$$v(\vec{x})\frac{\partial c}{\partial t}(\vec{x},t) = -\vec{\nabla}\cdot\vec{f}(\vec{x})c(\vec{x},t) \tag{5.1}$$

Here $0 < v < 1$ (dimensionless) is the blood volume fraction, and $\vec{f}$ in units of $mL/min/cm^2$ is the flow of blood (mL/min) through a unit tissue area (cm^2). The local tissue concentration c ($mmol/mL$) measures the amount of CA ($mmol$)

[1] 10.5281/zenodo.10056112 or https://github.com/EShalom/tkspace

per mL of blood, and is not directly measurable. Rather, what is measured in MRI is the tissue concentration C ($mmol/mL$), or the amount of CA per mL of tissue:

$$C(\vec{x}, t) = v(\vec{x})c(\vec{x}, t) \tag{5.2}$$

It is assumed throughout that blood and tissue are incompressible, so that $\vec{f}$ is divergence free:

$$\vec{\nabla} \cdot \vec{f}(\vec{x}) = 0 \tag{5.3}$$

This model therefore has 3 free parameters per interior voxel, with 4 scalar fields $(v, \vec{f})$, and one degree of freedom removed by the flow incompressibility. Expressing the transport equation in terms of the tissue concentration shows this more explicitly:

$$\frac{\partial C}{\partial t} = -\vec{\nabla} \cdot \vec{u}C \tag{5.4}$$

Here the blood velocity $\vec{u} = \vec{f}/v$ is introduced, which represents 3 degrees of freedom and is *not* divergence-free unless the blood volume fraction v is constant.

Without further constraints, the solution for $(v, \vec{f})$ can only be determined up to a constant: if $(v, \vec{f})$ solve the transport equations, then $(\alpha v, \alpha \vec{f})$ are solutions too, for any constant α. Therefore an additional constraint is needed to pin down the values unambiguously, for instance by assuming the volume fraction in one particular reference location is known. One approach could be to ensure that the spatial resolution is sufficiently high so that some voxels can be found which lie entirely inside a venous vessel. If these voxels are at location $\vec{x}_0$, one can safely assume that $v(\vec{x}_0) = 1$. Alternatively, if the velocity $\vec{u}$ provides sufficient information for the particular application, and volumes or blood flows are not required, the model can be solved in the tissue concentration picture directly (Eq. 5.4).

Two-compartment perfusion model

The two-compartment perfusion model involves an arterial (a) and venous (v) blood compartment, with mono-directional exchange from a to v by a perfusion field $F(\vec{x})$. Dropping coordinates from the definitions from here on for simplicity,

the system is defined by:

$$v^a \frac{\partial c^a}{\partial t} = -\vec{\nabla} \cdot \vec{f}^a c^a - F c^a \tag{5.5}$$

$$v^v \frac{\partial c^v}{\partial t} = -\vec{\nabla} \cdot \vec{f}^v c^v + F c^a \tag{5.6}$$

Where superscript indicates compartmental transport coefficients, concentrations, and volume fractions. The total volume fraction in these systems is constrained as $0 \leq v^a + v^v \leq 1$. Since the total blood flow is incompressible, this results in:

$$\vec{\nabla} \cdot (\vec{f}^a + \vec{f}^v) = 0 \tag{5.7}$$

Incompressibility of blood in the arterial compartment entails that the perfusion field (F) can be derived from the arterial flow (f^a):

$$F = -\vec{\nabla} \cdot \vec{f}^a \tag{5.8}$$

The model has 7 free parameters per interior voxel: 8 scalar fields ($v^a, \vec{f}^a; v^v, \vec{f}^v$), with one degree of freedom again removed via the flow incompressibility. The measurable quantity is the total tissue concentration:

$$C = v^a c^a + v^v c^v \tag{5.9}$$

The transport equations can be written in terms if the tissue concentrations $C^a = v^a c^a$ and $C^v = v^v c^v$ by defining arterial- and venous velocities $\vec{u}^a = \vec{f}^a / v^a$ and $\vec{u}^v = \vec{f}^v / v^v$ and the perfusion rate constant $K^{va} = F / v^a$:

$$\frac{\partial C^a}{\partial t} = -\vec{\nabla} \cdot \vec{u}^a C^a - K^{va} C^a \tag{5.10}$$

$$\frac{\partial C^v}{\partial t} = -\vec{\nabla} \cdot \vec{u}^v C^v + K^{va} C^a \tag{5.11}$$

This representation expresses the models directly in terms of 7 unconstrained scalar fields. As for the one-compartment case, the volumes and flows are only

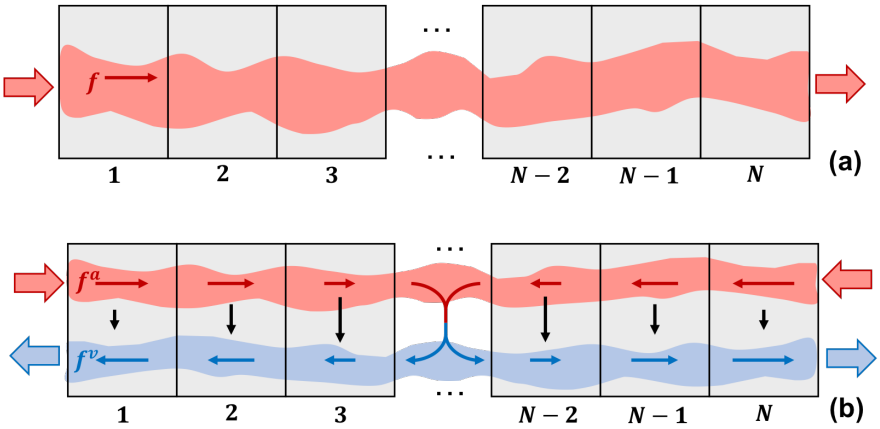

Figure 5.1: Illustration of the discretised compartment models. (A) A one-compartment blood flow model, showing a system with a positive flow direction (left-to-right). (B) A two-compartment perfusion model with arterial influxes and venous outfluxes at either end.

determined up to a constant: if $(v^a, \vec{f}^a)$ and $(v^b, \vec{f}^b)$ solve the equations, then $(\alpha v^a, \alpha \vec{f}^a)$ and $(\alpha v^b, \alpha \vec{f}^b)$ are solutions too, for any constant α. As before, the solution can be pinned down by adding a constraint such as $v^v(\vec{x}_0) = 1$ for some suitably chosen location $\vec{x}_0$ in a large venous vessel.

5.2.2 Discrete one-dimensions systems

Since the delivery of nutrients to tissue is a function of blood flow rather than blood velocity, clinical utility most likely hinges on the ability to measure flow. Hence all systems are simulated using the $(v, \vec{f})$ representation. In order to apply the spatiotemporal compartment models to a system of N voxels measured at K time points, an upwind discretisation is applied in space and first-order discretisation in time. The result is illustrated in figure 5.1.

One-compartment blood flow model

After discretisation, the one-compartment spatiotemporal model reduces to an N-compartment temporal model Sourbron (2014):

$$c_i(t + \Delta t) = c_i(t) + \frac{\Delta t}{v_i} \{k_{i,i-1} c_{i-1}(t) + k_{i,i+1} c_{i+1}(t) - k_i c_i(t)\} \tag{5.12}$$

Where Δt denotes the timestep, and quantities $c_i(t)$ and v_i defined at the voxel centre. The rate constants k_{ij} from j to i are positive and defined by:

$$k_{i,i-1} = \begin{cases} \frac{f_i}{\Delta x}, & f_i > 0 \\ 0, & f_i \leq 0 \end{cases} \qquad k_{i-1,i} = \begin{cases} 0, & f_i \geq 0 \\ \frac{-f_i}{\Delta x}, & f_i < 0 \end{cases} \tag{5.13}$$

Here the flow f_i is defined at the left interface of voxel i, and $k_i = k_{i-1,i} + k_{i+1,i}$. Additional free parameters to the model are the concentrations $c_0(t)$ and $c_{N+1}(t)$ at the left and right boundary of the system, respectively. System influxes ($J_+(t)$ and $J_-(t)$) are defined using these boundary concentrations ($c_0(t)$ and $c_{N+1}(t)$) with the corresponding rate constants. In a 1D scenario, the incompressibility of flow implies that it is constant: $f_i = f$. Hence the 1D one-compartment model is fully defined by the $1 + N$ quantities (f, v_i). For numerical stability, the time step Δt must be chosen to be smaller than the smallest voxel mean transit time:

$$\Delta t < \min_i \left\{ \frac{v_i}{k_i} \right\} \tag{5.14}$$

Two-compartment perfusion model

After discretisation, the two-compartment spatiotemporal model becomes a system of 2N temporal compartments:

$$c_i^a(t + \Delta t) = c_i^a(t) + \frac{\Delta t}{v_i^a} \left\{ k_{i,i-1}^a c_{i-1}^a(t) + k_{i,i+1}^a c_{i+1}^a(t) - (k_i^a + F_i) c_i^a(t) \right\} \tag{5.15}$$

$$c_i^v(t + \Delta t) = c_i^v(t) + \frac{\Delta t}{v_i^v} \left\{ k_{i,i-1}^v c_{i-1}^v(t) + k_{i,i+1}^v c_{i+1}^v(t) - k_i^v c_i^v(t) + F_i c_i^a(t) \right\} \tag{5.16}$$

Additional free parameters to the model are the arterial- and venous concentrations $c_0^a(t), c_0^v(t)$ and $c_{N+1}^a(t), c_{N+1}^v(t)$ at the left and right boundary of the system, respectively. System influxes ($J_+^a(t)$ and $J_-^a(t)$) are defined using the arterial boundary concentrations ($c_0^a(t)$ and $c_{N+1}^a(t)$) with the corresponding rate constants. In 1D systems, the incompressibility of the flow implies that the total flow is constant ($f_i^a + f_i^v = f$) and that the arterial flow at the right

boundary of a voxel is that at the left boundary minus the loss by perfusion:

$$f^a_{i+1} = f^a_i - F_i \Delta x \tag{5.17}$$

This implies that the arterial flow at any boundary i is fully defined by the field F_i and the arterial flow f^a_0 at the left boundary. For given total flow f the venous flow is then also determined everywhere ($f^v_i = f - f^a_i$). Hence in the flow picture the discrete system is fully defined by the $3N + 2$ quantities $(v^a_i, v^v_i, F_i, f^a_0, f^v_0)$. For numerical stability, the time step must be smaller than the smallest voxel mean transit time:

$$\Delta t < \min_i \left\{ \frac{v^a_i}{k^a_i + F_i}, \frac{v^v_i}{k^v_i} \right\} \tag{5.18}$$

5.2.3 Parameter reconstruction

The measured data consist of a 2D tissue concentration matrix C^{meas}_{ik} with one value for each voxel i and each time point k. For given values of the discrete model parameters (volume fractions, flows and boundary concentrations), a predicted concentration C^{pred}_{ik} is generated by iterating the discrete equations (5.12 or 5.15, 5.16) with a time step Δt satisfying Eqs. (6.17, 5.18). The resulting concentrations at high temporal resolution are then downsampled to the measured temporal resolution and scaled with the volume fractions (Eq. 5.2 or 5.9).

Optimal values for the model parameters are determined by minimising the root-mean-square difference between C^{meas}_{ik} and C^{pred}_{ik}. The initial guesses for the total volume fraction (v) and the boundary concentrations are estimated from the data. The unknown boundary concentrations are estimated from the concentrations at the voxel nearest to the boundary. The volume fraction, up to a scaling constant, is estimated from the concentration at the last time point. Assuming state-state has been reach at this time, tissue concentrations are directly proportional to v.

The optimisation is performed iteratively over time: parameters are first optimised using only data up to an initial time t_0 - chosen to be after the initial peak of concentration has entered the system; subsequently the next time point

is added and the parameters are optimised again, using the solutions from the previous step as initial values. This process is repeated until all time points are added.

The optimisation for each time step is performed by a second-order gradient descent, after normalising the parameters to dimensionless quantities in the range $[0, 1]$. For parameter values at the lower or upper bounds, a first-order method is applied. For flow values close to zero, the gradient is evaluated at zero. For a given gradient, the parameters are updated using an Adams update based linesearch (Kingma & Ba, 2015). In this Adams update based approach, the moving averages of gradient and squared gradient are used to guide the optimisation. The resulting update is scaled to restrict parameter updates crossing zero.

The algorithm is implemented in python and can be found on GitHub.[1]

5.2.4 Simulations

The parameter reconstruction was evaluated for the one-compartment blood flow model and the two-compartment perfusion model. For each model, three digital reference objects were evaluated, detailed in Table 5.1 and Table 5.2. Models have total spatial dimensions of 25.6cm with Δx=0.8cm, evolved to a total time of 80s. Uniqueness and sensitivity of the solution were estimated by repeating the reconstruction with different initial guesses, noise levels and levels of temporal undersampling. Reconstruction accuracy was measured for each parameter field P by the difference between reconstruction (P_{rec}) and ground truth (P_{gt}) as a percentage the mean absolute parameter value:

$$E_{rel}(P) = \left| \frac{P_{rec} - P_{gt}}{\mu(|P_{gt}|)} \right| \times 100. \tag{5.19}$$

For comparison, the mean and standard deviation of the resulting $E_{rel}(P)$ distributions are reported denoted by $\bar{E}_{rel}$.

Since volume fractions and flows are only determined up to a constant, reconstruction accuracy is measured only for the velocities and rate constants, which are not subject to this redundancy.

[1] 10.5281/zenodo.10056112 or https://github.com/EShalom/tkspace

Table 5.1: Ground truth values for the one-compartment systems. All x values used are in cm. P_{AIF} is a population AIF (Parker *et al.*, 2006) with a defined delay (d) and a scaling factor ($0 \leq s_f \leq 1$)

	Ground Truth Case		
Parameter	**1**	**2**	**3**
f (ml/s/cm^2)	1	0.5	-0.6
v (ml/ml)	$0.24\sin^2(0.3x)$ $+0.36\cos^2(0.15x)+0.3$	$0.3\sin^2(0.1x)$ $+0.36\cos^2(0.2x)+0.3$	$0.24\sin^2(0.3x)$ $+0.36\cos^2(0.15x)+0.3$
J_+ (mM/s)	$P_{AIF}(d=10s, s_f=1)$	$P_{AIF}(d=15s, s_f=1)$	0
J_- (mM/s)	0	0	$P_{AIF}(d=15s, s_f=1)$

Reconstructions of noiseless data with 2 second temporal resolutions were repeated for several sets of initial values. For the one-compartment cases, these were f=$\pm(11, 9, 7, 4, 3, 1)$. For the two-compartment systems the initial value sets are detailed in Table 5.3.

Sensitivity to temporal undersampling was tested by reconstructing the noiseless systems with data sampled at 2, 4, 6, 8, and 10s. Sensitivity to noise was tested by repeating reconstructions on data with signal-to-noise ratio (SNR) levels of 5, 10, 15, and 20. Gaussian noise was added with a standard deviation (σ) derived from the mean concentration:

$$\sigma = \frac{\bar{C}^{meas}}{SNR} \tag{5.20}$$

The SNR lower limit of 5 was chosen to reflect the typical lower limit used in DCE-MRI protocols (Banerji *et al.*, 2012). For each SNR level, reconstructions are run with a given set of initial values for 5 realisations to calculate 95% confidence intervals on the reconstructed parameters.

Computations are run on a single CPU (Intel(R) Xeon(R) Gold 6152 CPU 2.10GHz), with a maximum of 10,000 iterations at each time iteration, and a gradient evaluation step 1×10^{-4} for one-compartment systems and 5×10^{-5} for two-compartment systems.

5.2.5 Sensitivity Analysis

To deduce the sensitivity of the two-compartment forward model to changes in the model parameters a simple univariate based analysis was carried out. This

Table 5.2: Ground truth values for the two-compartment systems. All x values used are in *cm*. $P_{AIF}(d, s_f)$ is a population-based AIF (Parker *et al.*, 2006) with a defined delay (d) and a scaling factor ($0 \leq s_f \leq 1$); $G(w, h)$ denotes a centred Gaussian with width (w) and height (h); and $Q(a, b, e)$ denotes a quadratic starting at a passing b at system centre and ending at e.

	Ground Truth Case		
Parameter	**1**	**2**	**3**
f_1^a (ml/s/cm^2)	0.9	0.512	0.3
f_1^v (ml/s/cm^2)	-0.5	-0.512	-0.6
F (ml/s/ml)	$G(0.5L_x, 0.0626)$	$G(0.16\dot{6}L_x, 0.1)$	$0.0336\sin^2((20000 - 6266.25\pi)x) + 0.021$
v (ml/ml)	$0.24\sin^2(0.3x) + 0.36\cos^2(0.15x) + 0.3$	$v^a + v^v$	$0.3\sin^2(0.2x) + 0.36\cos^2(0.3x) + 0.3$
λ^a	$0.3\cos^2(0.01x) + 0.3$	v^a/v	$0.3\sin^2((0.007x) + 0.3$
v^a (ml/ml)	$\lambda^a v$	$\lvert f^a\rvert/Q(19, 4.9, 19)$	$\lambda^a v$
v^v (ml/ml)	$v - v^a$	$\lvert f^v\rvert/Q(7.1, 1.5, 7.1)$	$v - v^a$
J_+^a (mM/s)	$P_{AIF}(10s, 0.6)$	$P_{AIF}(10s, 1)$	$P_{AIF}(10s, 1)$
J_-^a (mM/s)	$P_{AIF}(15s, 0.4)$	$P_{AIF}(10s, 1)$	$P_{AIF}(15s, 1)$

Table 5.3: Initial guesses for parameters applied over all two-compartment cases.

	Guess Set				
Parameter	**1**	**2**	**3**	**4**	**5**
f_1^a (ml/s/cm^2)	1.2	1.2	0.5	0.8	0.3
f_1^v (ml/s/cm^2)	-1.2	-0.3	-1.0	-0.8	-1.2
F (ml/s/ml)	0.065	0.055	0.045	0.035	0.025
λ^a	0.4	0.45	0.5	0.55	0.6

analysis adjusted each parameter (3N+2+2T total) by 5% of the ground truth value in each case, with the impact on the resulting total tissue concentrations calculated relative to the maximum system tissue concentration (Equation 5.21).

$$\Delta C_{rel}^{i,k} = \frac{(C_{gt}^{i,k} - C_{adjusted}^{i,k})}{\max(C_{gt})} \tag{5.21}$$

Here, i and k denote each spatial and temporal point, respectively. The differences are scaled as a percentage of the tissue concentration maximum for comparison between cases.

To deduce the covariance ($\mathfrak{X}$) of the two-compartment model the approach from Scipy `curve_fit` (Virtanen *et al.*, 2020) was used. This applies a Moore-Penrose pseudoinverse (Moore, 1920; Penrose, 1955) to compute the covariance matrix from the Hessian matrix, $\mathbf{H}$, comprising of the second derivatives of the object function and the model parameters. The covariance is defined as standard in a least squares problems as the matrix inverse of the Hessian matrix (Gejadze *et al.*, 2008).

$$\mathfrak{X} = \mathbf{H}^{-1} \tag{5.22}$$

The Moore-Penrose pseudoinverse method to caluclate this requires the formation of a Jacobian matrix ($\mathfrak{J}$) with terms for the gradient of the object function at every coordinate point with respect to each individual parameters. The Jacobian matrix constructed therefore has NT rows for each total tissue concentration coordinate value and $3N + 2 + 2T$ columns for the model parameters ($f_0^a, f_0^v, F(x), v^a(x), v^v(x), C_+^a(t), C_-^a(t)$). The Jacobian matrix takes the form:

$$\mathfrak{J} = \begin{bmatrix} \frac{\partial C_0}{\partial P_0} & \cdots & \frac{\partial C_m}{\partial P_0} \\ \vdots & \ddots & \vdots \\ \frac{\partial C_0}{\partial P_n} & \cdots & \frac{\partial C_m}{\partial P_n} \end{bmatrix} \tag{5.23}$$

The singular value deposition (SVD) (Stewart, 1993) of the resulting Jacobian matrix was then calculated using `scipy.linalg.svd` which factorises

$\mathfrak{J}$ into unitary matrices U, V and a diagonal matrix S defined by:

$$\mathfrak{J} = USV^T; \mathfrak{J}^T = VS^TU^T \tag{5.24}$$

Where the pseudoinverse of the Jacobian matrix and its transpose are defined using the results from the SVD via:

$$\mathfrak{J}^{-1} = VS^{-1}U^T; (\mathfrak{J}^T)^{-1} = U(S^T)^{-1}V^T \tag{5.25}$$

The Jacobian matrix can be used to provide an estimate for the Hessian matrix by matrix multiplication with its transpose ($\mathfrak{J}^T$). Using the definitions of $\mathfrak{J}$ and $\mathfrak{J}^T$ above the Hessian can then be estimated directly from the SVD matrices (Strang, 1980):

$$\mathbf{H}^{-1} \approx (\mathfrak{J}^T\mathfrak{J})^{-1} = \mathfrak{J}^{-1}(\mathfrak{J}^T)^{-1} = VS^{-1}S^{-1}V^T \tag{5.26}$$

Due to parameter magnitude differences the covariances are converted to correlation coefficients, ρ, using the product of the parameter standard deviations (Equation 5.27):

$$\rho = \mathfrak{D}^{-1}\mathfrak{X}\mathfrak{D}^{-1}; \quad \text{with} \quad \mathfrak{D} = diag(\sqrt{\mathfrak{X}}) \tag{5.27}$$

The resulting correlation coefficient matrix has a central diagonal of +1 as all parameters self correlate exactly, with correlation between different parameters shown by the off-diagonal elements.

5.3 Results

Results from both one-compartment and two-compartment systems are summarised within this section. Additional data tables of $\bar{E}_{rel}$ values are included in Appendix B.

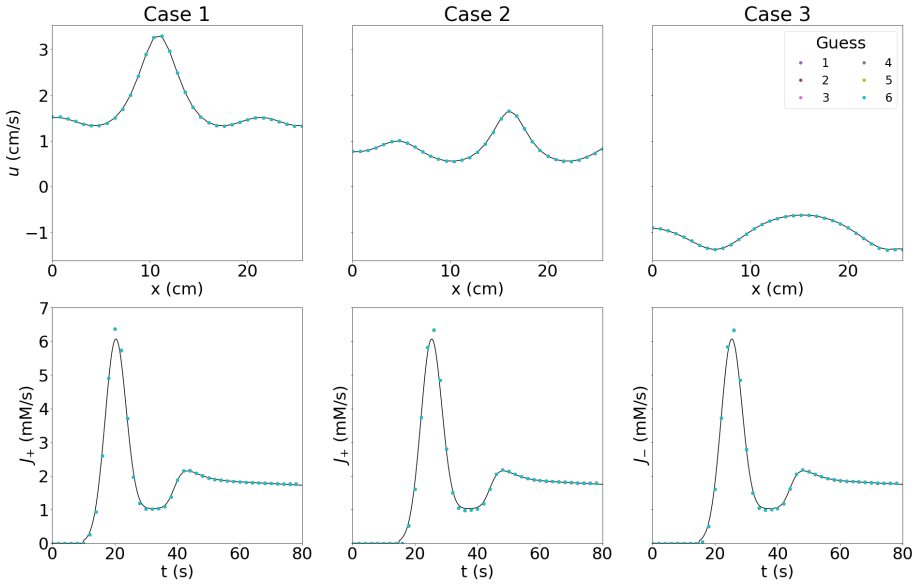

Figure 5.2: Parameter reconstructions for all noise-free one-compartment system cases. The solid black line indicates the ground truth, and the coloured dots show reconstructions with different initial guesses.

5.3.1 One-compartment blood flow model

Results for the noise-free one-compartment systems are summarised in Figure 5.2, showing the solutions are accurate and independent of the initial guesses. The average error $\bar{E}_{rel}$ (mean $\pm$ standard deviation) across all cases is $2.9 \pm 4.7\%$ and $0.4 \pm 0.3\%$ for J and u respectively.

Concentration-time data reconstructed from recovered parameter values at one of the initial guesses are shown in Figure 5.3. Deviations between the ground truth concentration and recovered parameter profiles are not visually detectable and are between $\pm 2.5\%$ of the maximal concentration in each case.

The effect of SNR and sampling interval on parameter reconstruction is summarised in Figure 5.4(a). Subject-level results are included in Figures 5.9-5.14. Average $\bar{E}_{rel}$ across all simulations at SNR 5 is $2.4 \pm 2.8\%$ and $16.1 \pm 15.6\%$ for u and J, respectively. Average $\bar{E}_{rel}$ across all simulations at Dt 8s is $14.4 \pm 15.3\%$ and $24.6 \pm 47.3\%$ for u and J, respectively. The results show the expected behaviour with increasing accuracy and precision at higher SNR and smaller Dt in all parameters. Velocities are substantially more robust to

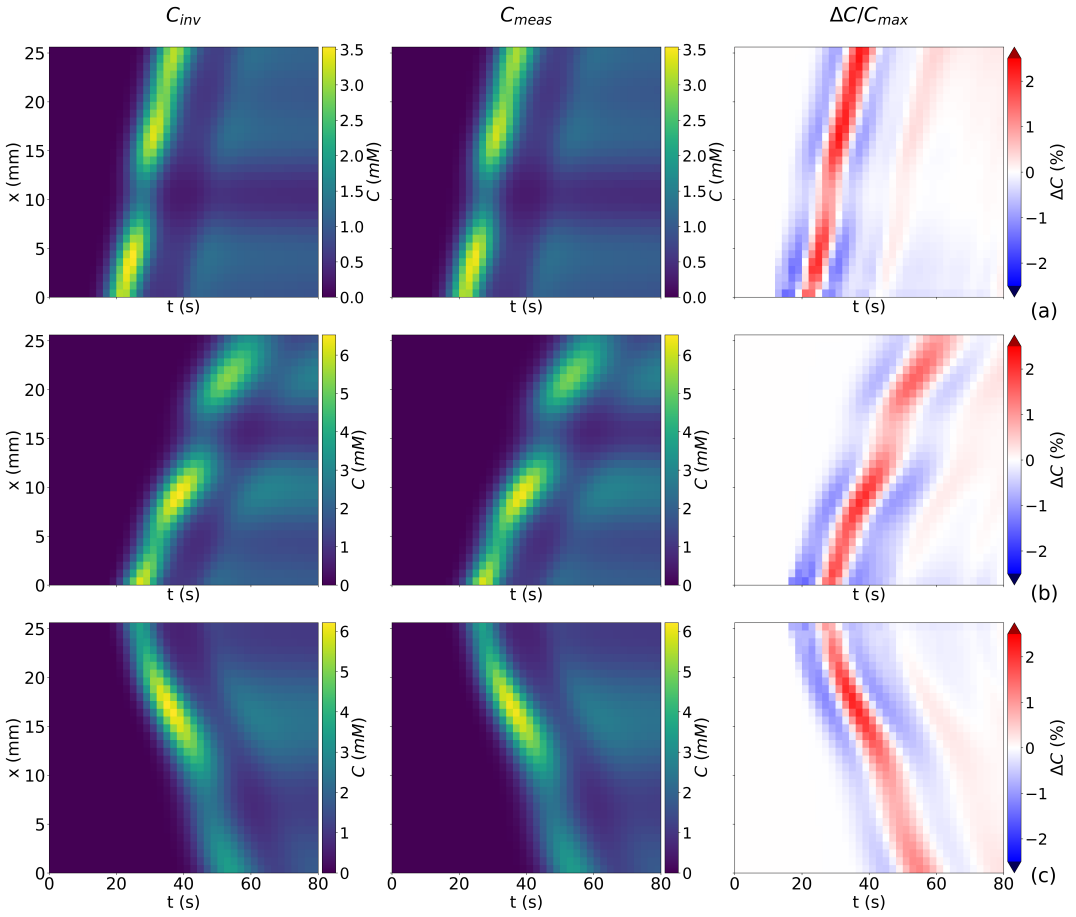

Figure 5.3: Comparison of the recovered concentration values from the retrieved parameters against the ground truth, alongside the percentage difference for the maximal concentration. Shown are cases 1, 2 and 3 in rows (a), (b), and (c), respectively. Differences above or below ±2.5% of the maximum concentration value are shown by dark red or dark blue, respectively.

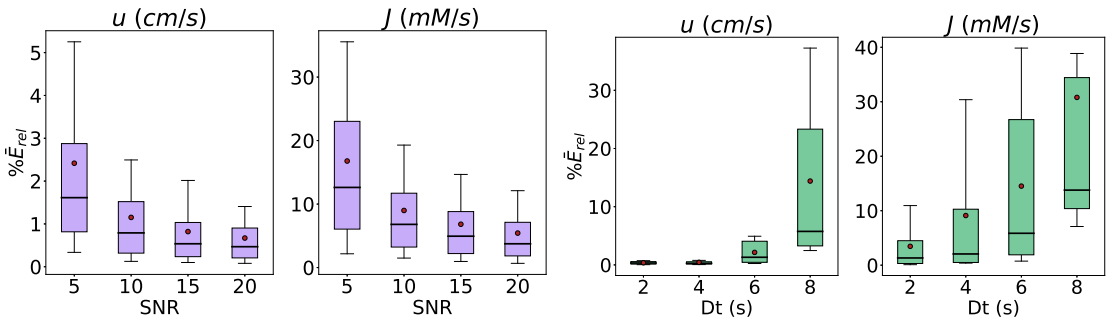

Figure 5.4: Box plots of reconstruction errors for all one-compartment parameters as a function of SNR (a) and temporal sampling Dt (b).

noise than the influxes, and more accurate and precise at smaller Dt levels.

5.3.2 Two-compartment perfusion model

Results for the noise-free two-compartment systems are summarised in Figure 5.5, showing that the solutions are generally less well determined than in the one-compartment case. Reconstruction of arterial velocity (u^a) and influx (J^a) is most accurate with lowest $\bar{E}_{rel}$ of $27.1 \pm 82.0\%$ and $13.0 \pm 38.1\%$, respectively. The perfusion rate (K^{va}) and venous velocity (u^v) are least precise with $\bar{E}_{rel}$ of $44.7\% \pm 76.5\%$, and $54.9\% \pm 121.4\%$ across all cases, respectively.

Ground-truth and reconstructed concentrations for initial guess 4 and noise-free data are shown in Figure 5.6. Despite the substantial errors in the reconstructed parameters, the reconstructed concentration is close to the ground truth and visually virtually indistinguishable. Since an accurate fit to the data is obtained with inaccurate parameters, this shows that multiple solutions are compatible with the observations.

Figures 5.7 and 5.8 shows the impact of SNR and undersampling on parameter accuracy and precision, showing the expected trend of increasing accuracy and precision at higher SNR and smaller Dt. The magnitude of the error is generally comparable between parameters except for the perfusion rate K^{va}, which is more sensitive to noise than the other parameters, and the venous velocity u^v, which appears particularly sensitive to undersampling. Figures 5.15- 5.20 illustrate these effects in more detail for all three example cases.

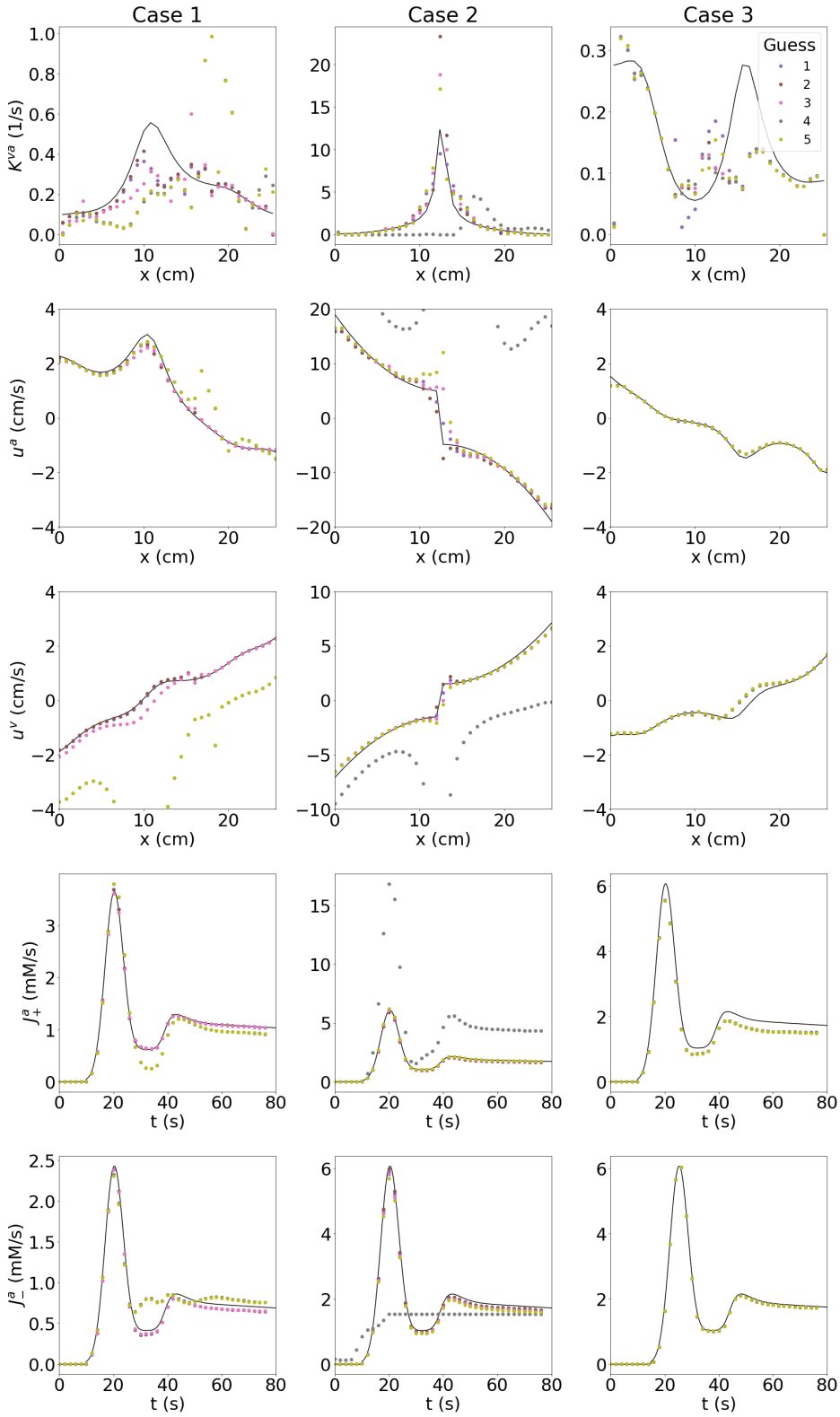

Figure 5.5: Parameter reconstructions for all noise-free two-compartment system cases. The solid black line indicates the ground truth, and the coloured dots show reconstructions with different initial guesses.

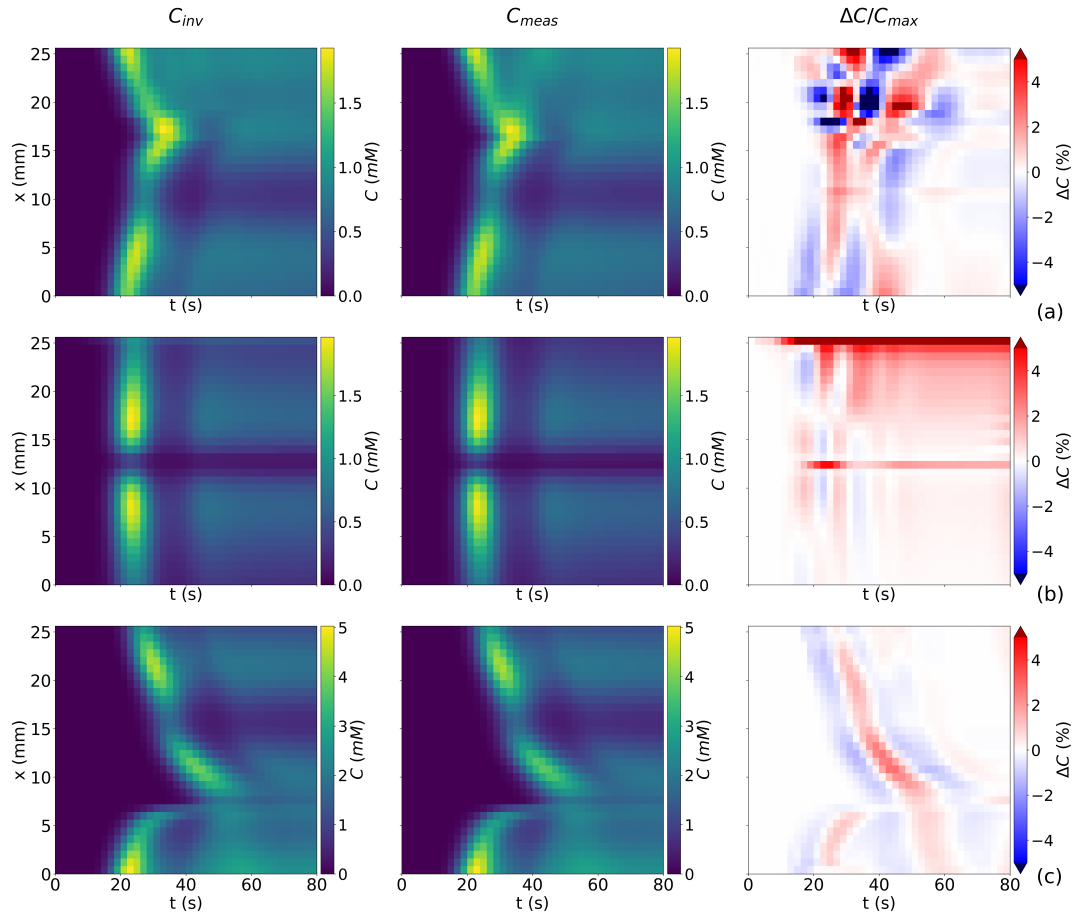

Figure 5.6: Comparison of the recovered concentration values from the retrieved parameters against the ground truth, alongside the percentage difference for the maximal concentration. Shown are cases 1, 2 and 3 in rows (a), (b), and (c), respectively for initial guess 4. Differences above or below ±5% of the maximum concentration value are shown by dark red or dark blue, respectively.

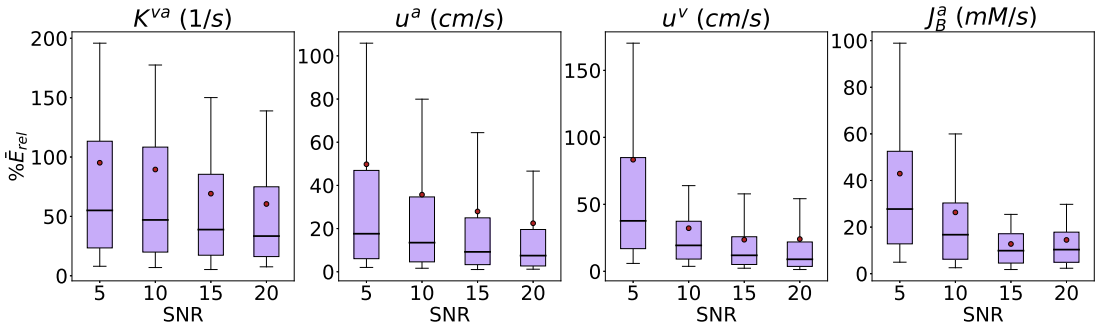

Figure 5.7: Distribution of each parameter error for the two-compartment system relative to the absolute mean parameter value within each system. The distribution is shown across all 3 cases for all noise realisations at each SNR.

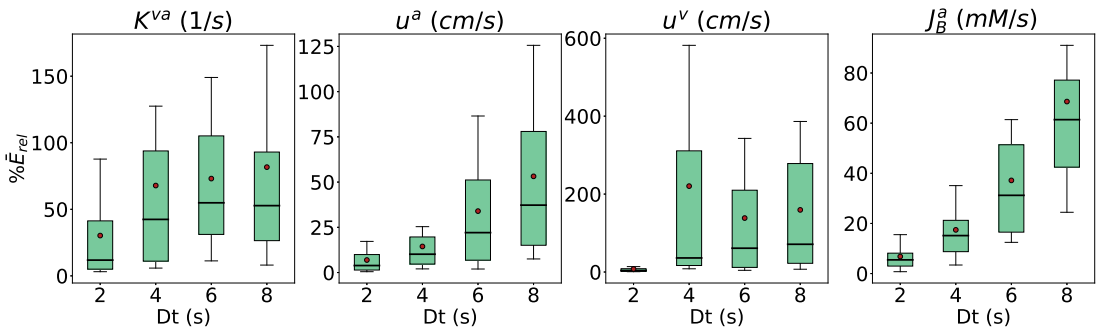

Figure 5.8: Distribution of each parameter error for the two-compartment system relative to the absolute mean parameter value within each system. The distribution is shown across all 3 cases for all undersampling rates.

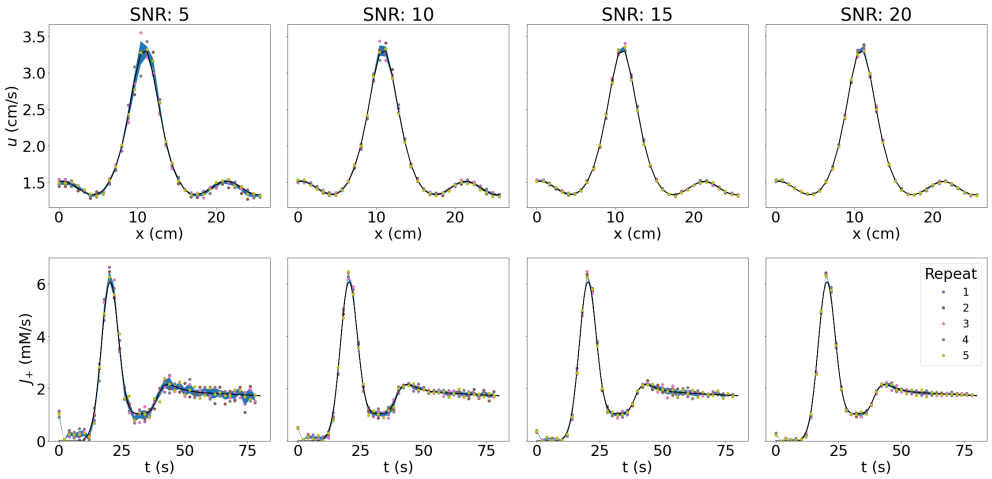

Figure 5.9: Parameter fitting results for SNR 5, 10, 15, 20 in one compartment for Case 1. The shaded area denotes the 95% confidence interval extracted from 5 noise realisation runs.

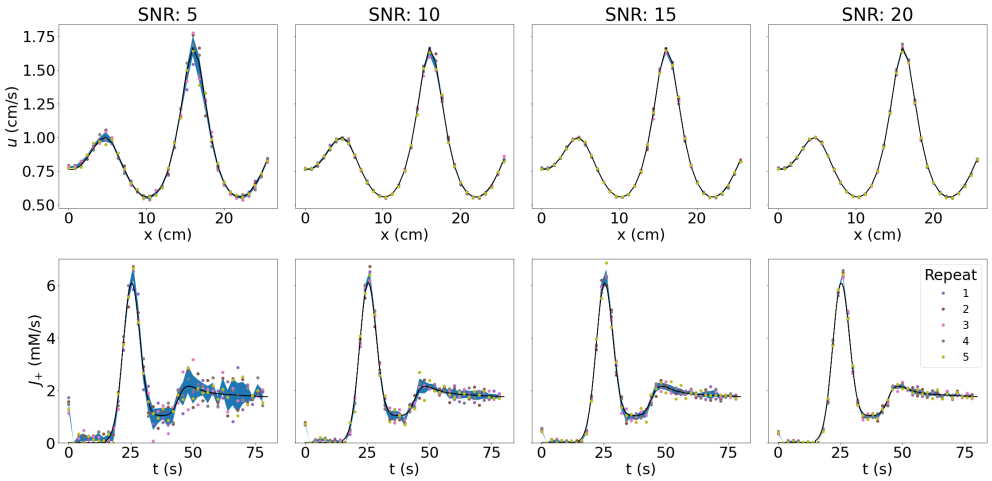

Figure 5.10: Parameter fitting results for SNR 5, 10, 15, 20 in one compartment for Case 2. The shaded area denotes the 95% confidence interval extracted from 5 noise realisation runs.

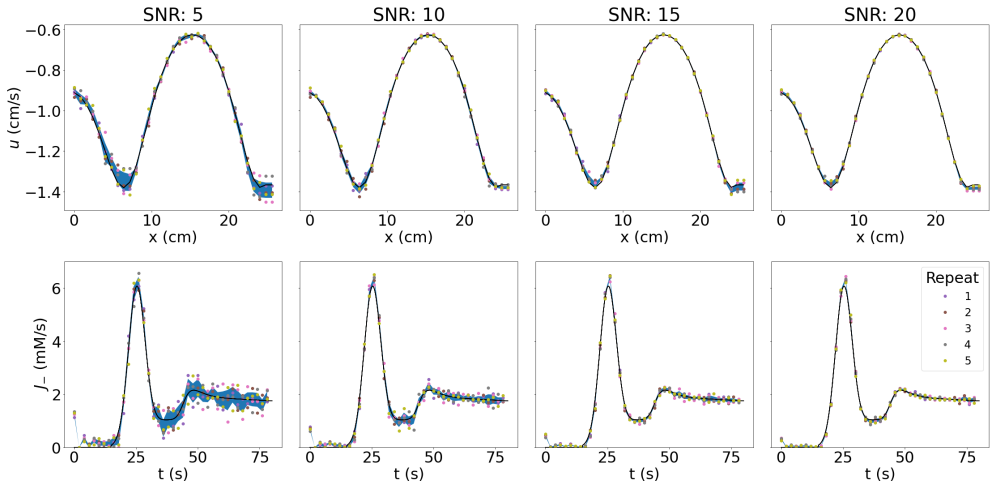

Figure 5.11: Parameter fitting results for SNR 5, 10, 15, 20 in one compartment for Case 3. The shaded area denotes the 95% confidence interval extracted from 5 noise realisation runs.

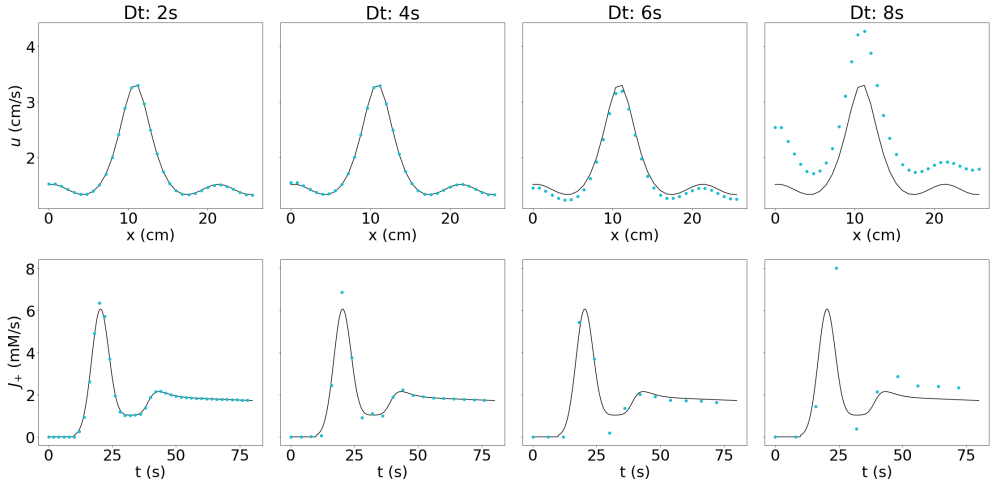

Figure 5.12: Parameter fitting results for time sampling resolution of 2, 4, 6, and 8s in one compartment for Case 1.

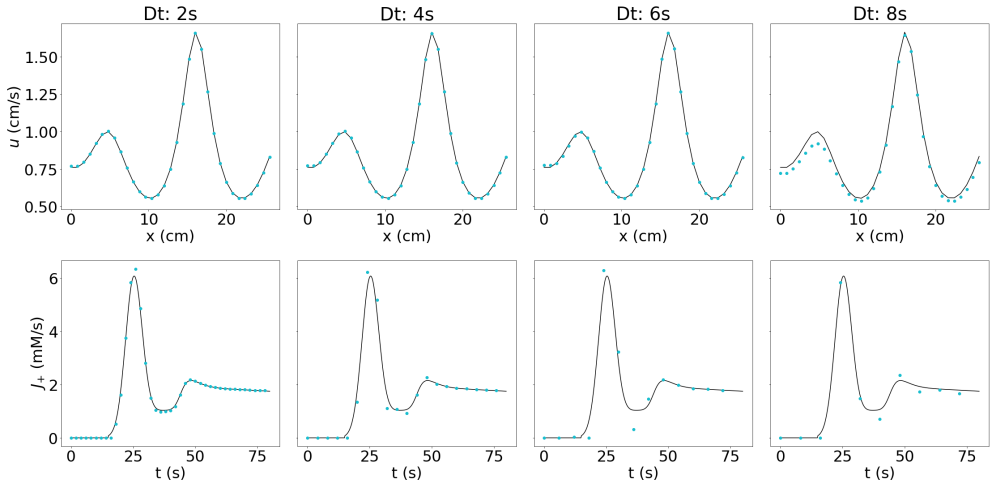

Figure 5.13: Parameter fitting results for time sampling resolution of 2, 4, 6, and 8s in one compartment for Case 2.

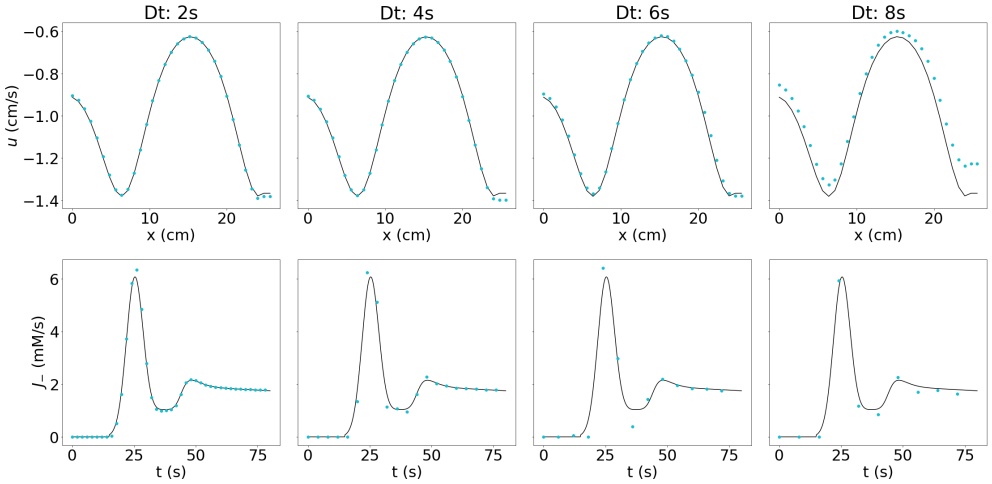

Figure 5.14: Parameter fitting results for time sampling resolution of 2, 4, 6, and 8s in one compartment for Case 3.

5.3.3 Sensitivity Analysis

Results for the two-compartment forward model sensitivity analysis are summarised in Figures 5.21 and 5.22. These illustrate the relative concentration difference distribution for the test cases. Parameters in the model are adjusted by a 5% increase or decrease from their ground truth values in Figures 5.21 and 5.22, respectively. Due to the large number of individual parameters Figures 5.21 and 5.22 report results by parameter group. All forward model differences for each individual voxel or time point value permutation are shown. For example, the arterial volume fraction (v^a) group distribution shown within Figure 5.21 consists of all $\Delta C_{rel}^{i,k}$ values for each v_i^a adjustment. Within Figures 5.21 and 5.22 the influence of adjusting individual parameters within the forwards model by ±5% produced similar $\Delta C_{rel}^{i,k}$, apart from the arterial boundary flow f_0^a which shows a much larger effect. The differences produced in the concentration-time data are largely indistinguishable between many of the individual parameters (Figure 5.23). Apart from f_0^a which shows $\Delta C_{rel}^{i,k}$ values ≈ 20%, other parameters show $\Delta C_{rel}^{i,k}$ values below 5% as illustrated for Case 1 (Figure 5.23).

The derived covariance analyses of the model parameters are illustrated by the correlation coefficients (Equation 5.27) for each of the 3 test cases (Figures 5.24-5.26). Each of the individual cases have slightly different trends in

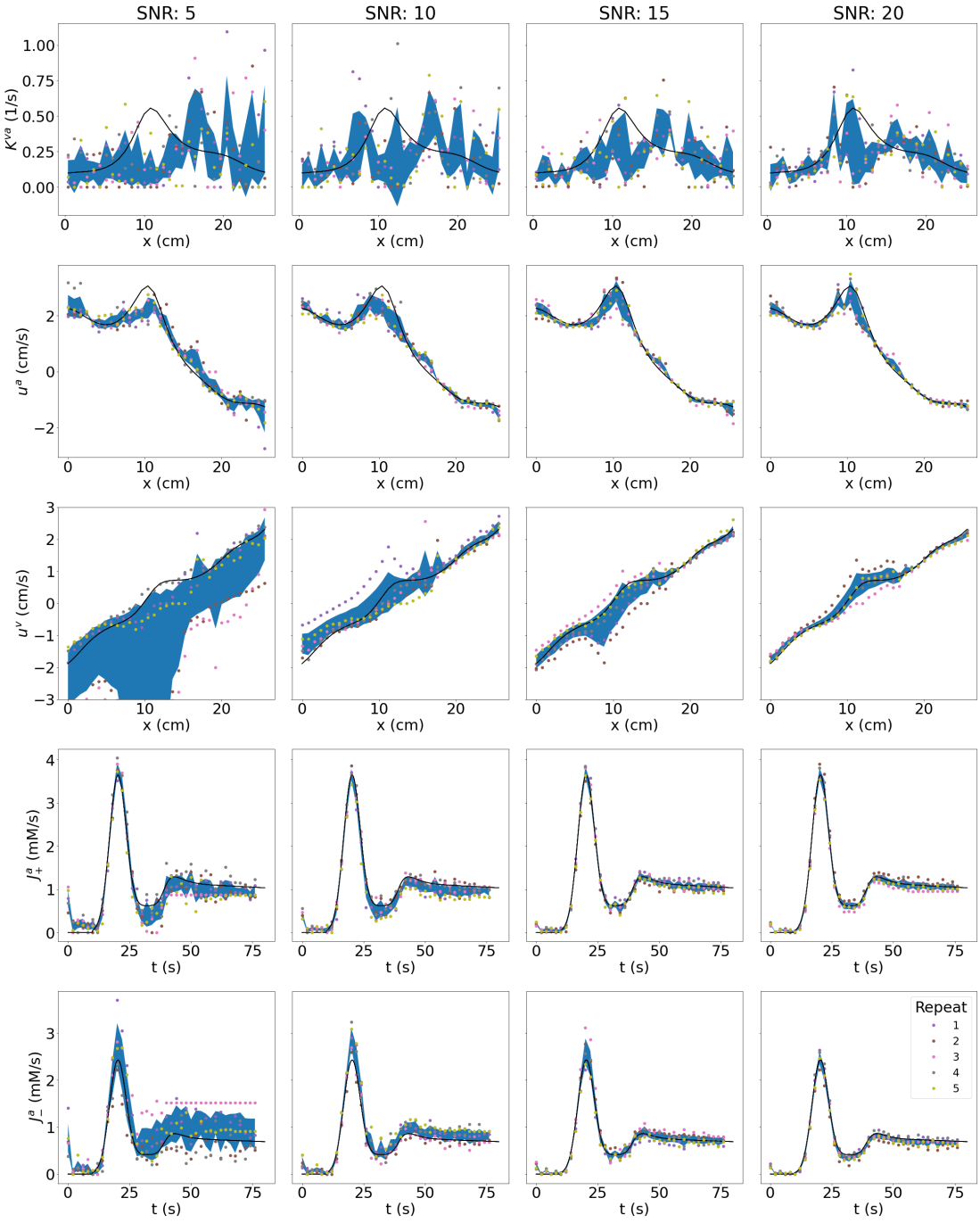

Figure 5.15: Parameter fitting results for SNR 5, 10 , 15, 20 in two-compartment Case 1. The shaded area denotes the 95% confidence interval extracted from 5 noise realisation runs.

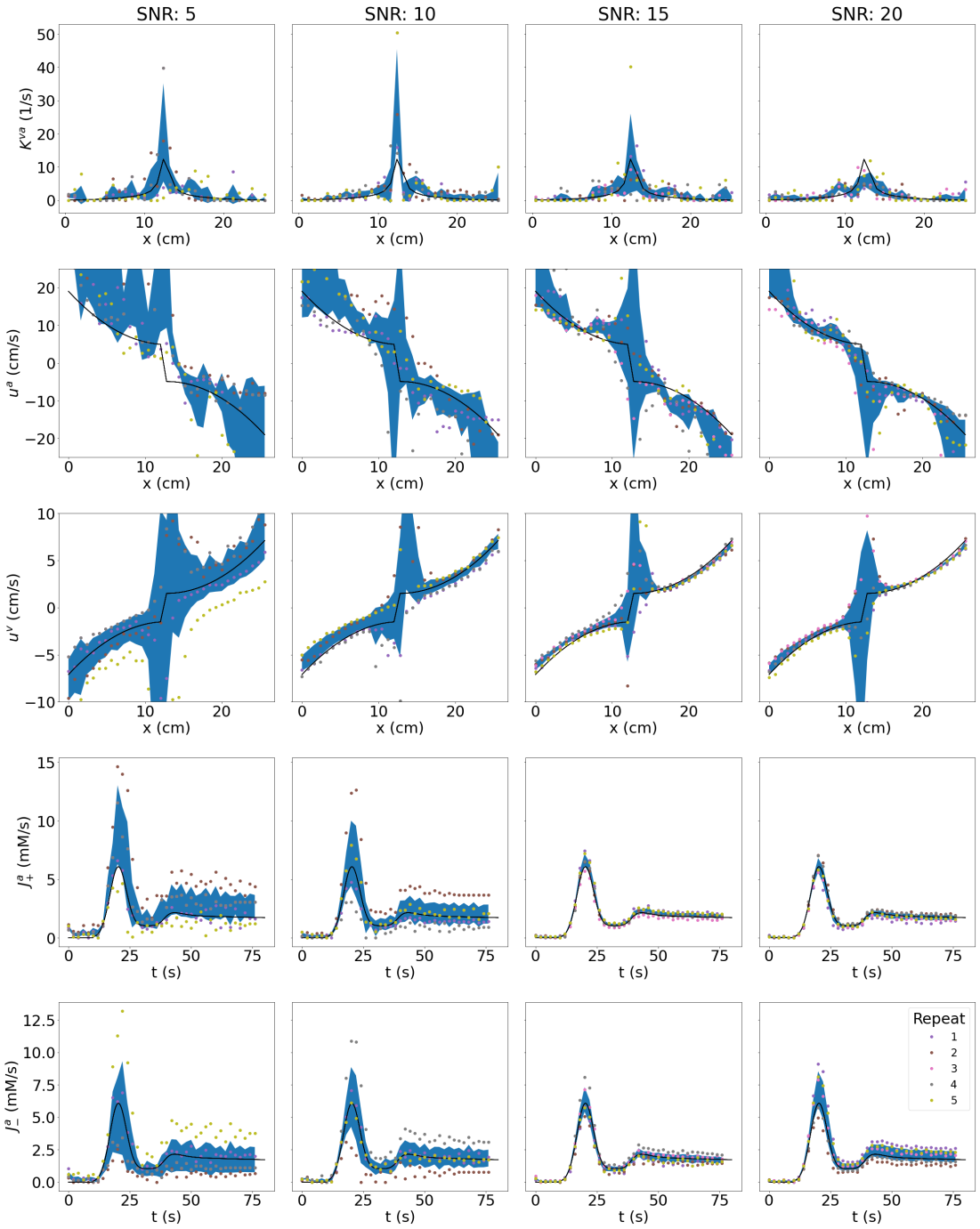

Figure 5.16: Parameter fitting results for SNR 5, 10, 15, 20 in two-compartment Case 2. The shaded area denotes the 95% confidence interval extracted from 5 noise realisation runs.

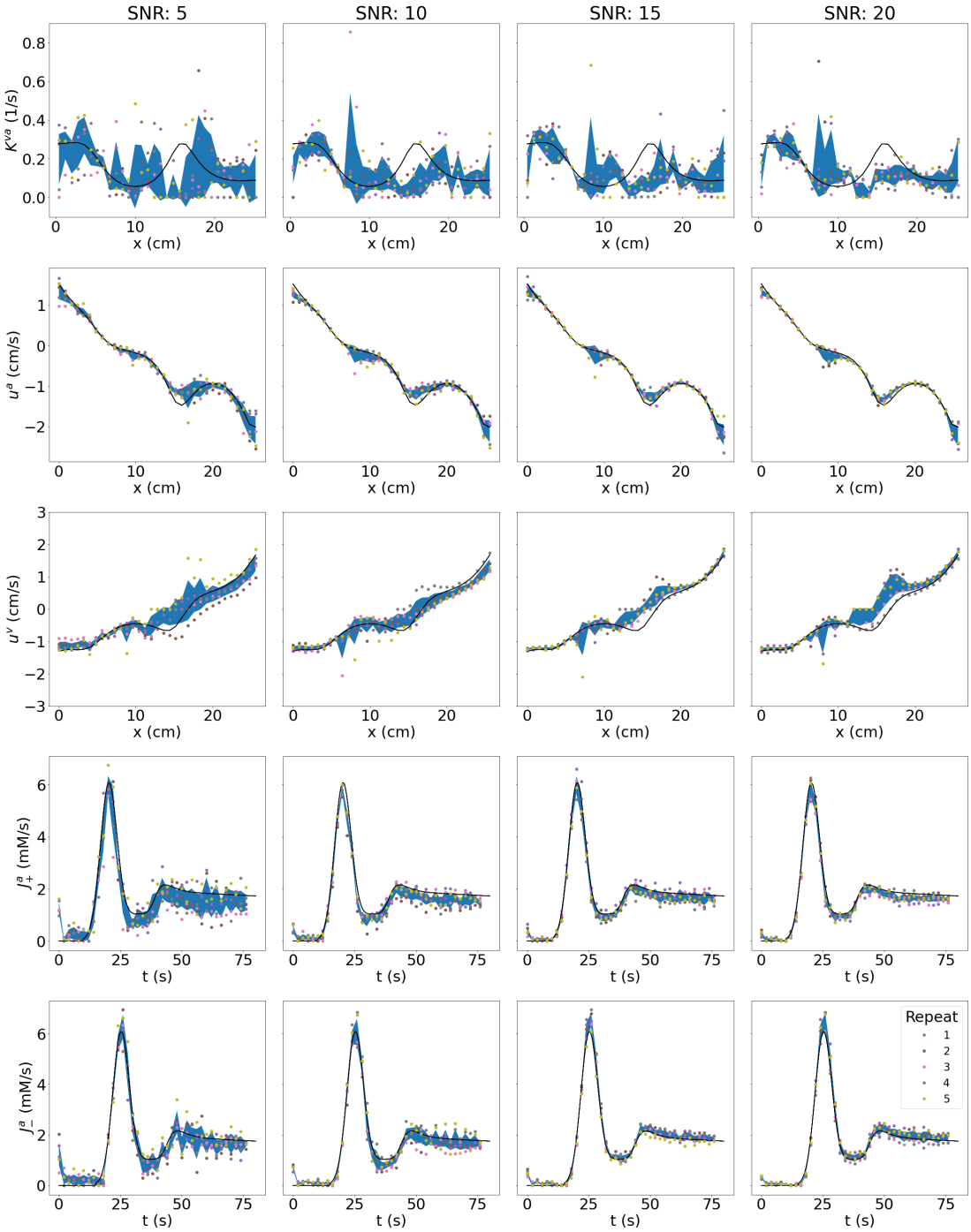

Figure 5.17: Parameter fitting results for SNR 5, 10 , 15, 20 in two-compartment Case 3. Here the black lines denote the ground truth parameter values with the shaded area showing the 95% confidence interval of parameter reconstructions extracted from the 5 noise realisation runs.

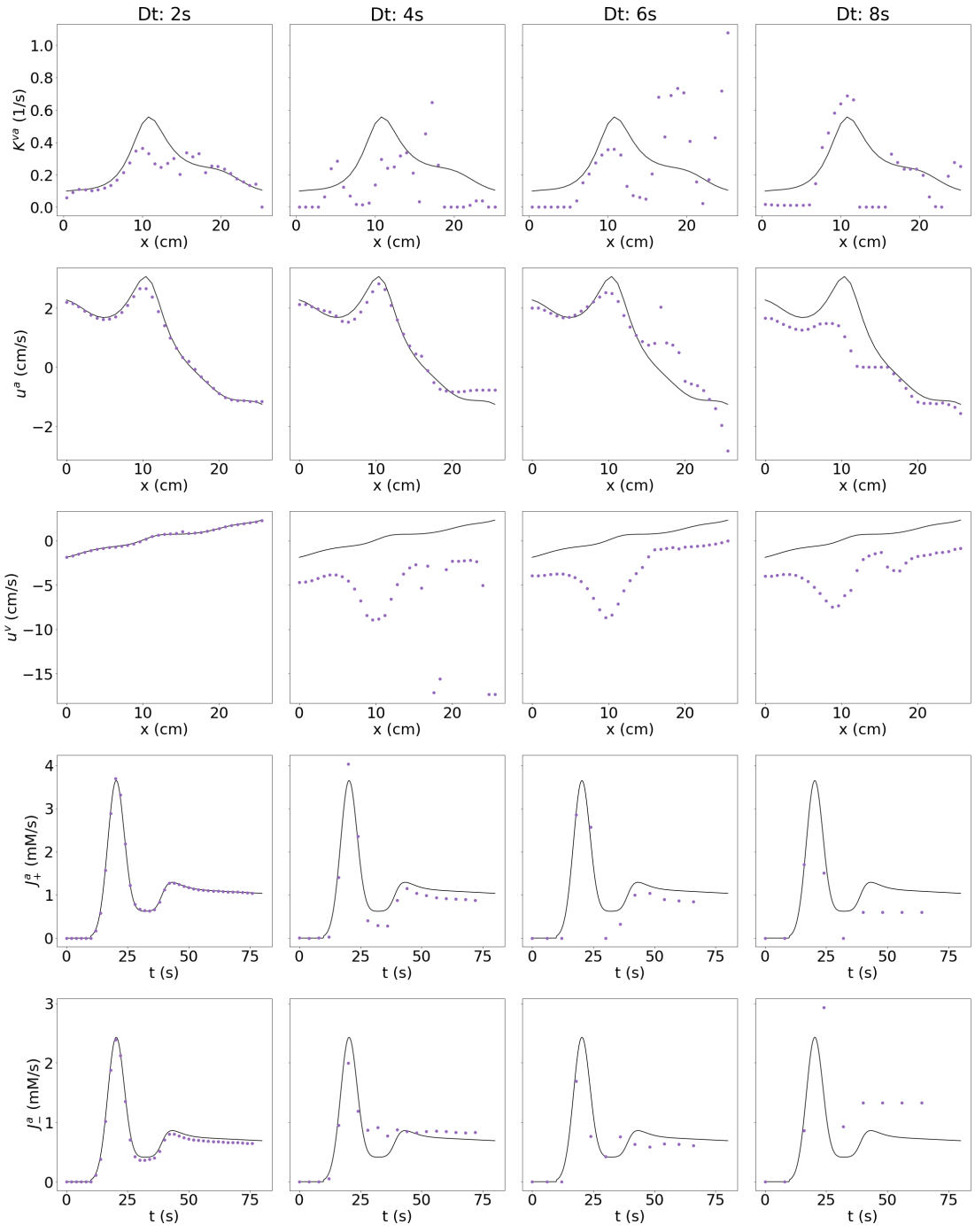

Figure 5.18: Parameter fitting results for time sampling resolution of 2, 4, 6 and 8s in two-compartment Case 1.

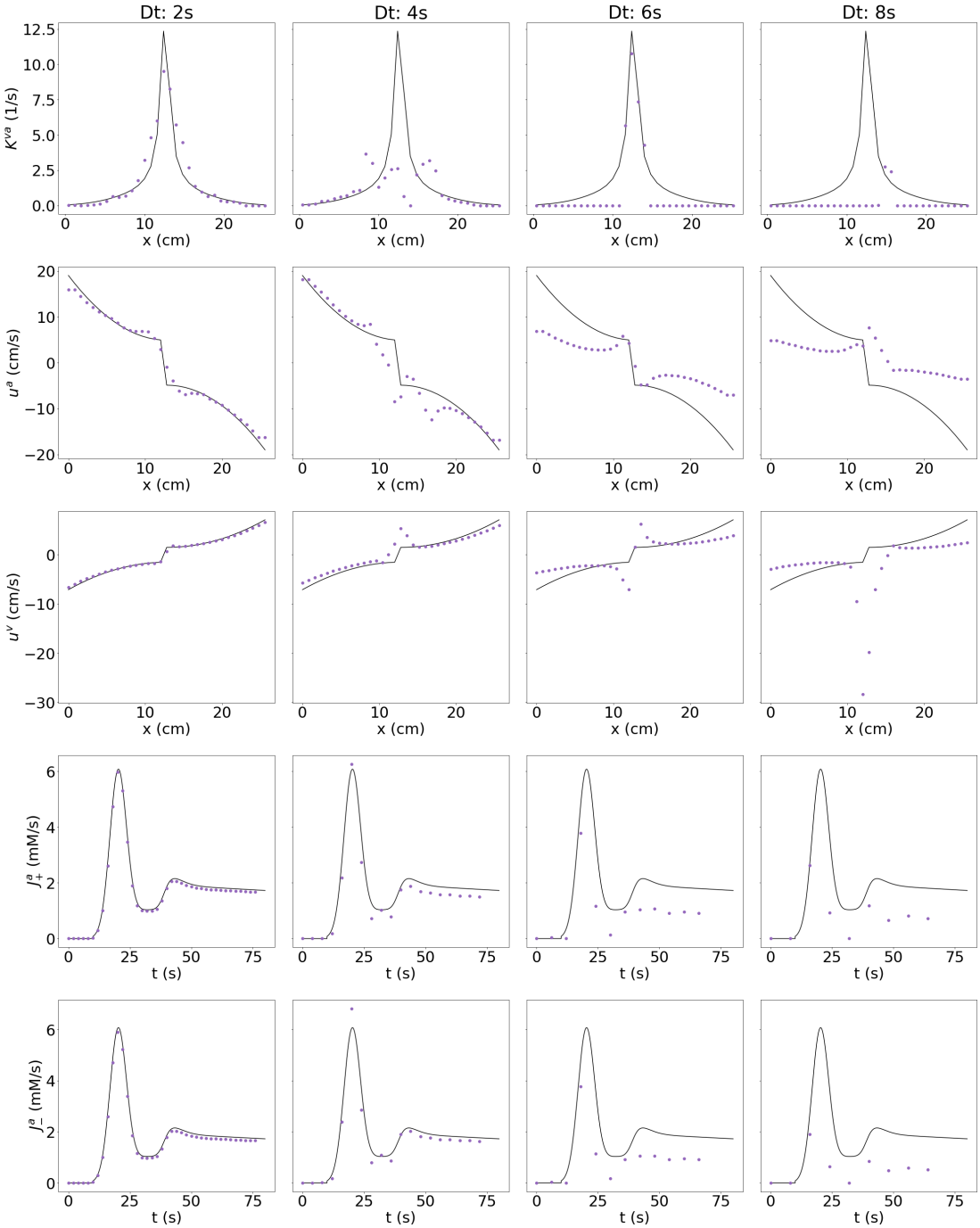

Figure 5.19: Parameter fitting results for time sampling resolution of 2, 4, 6 and 8s in two-compartment Case 2.

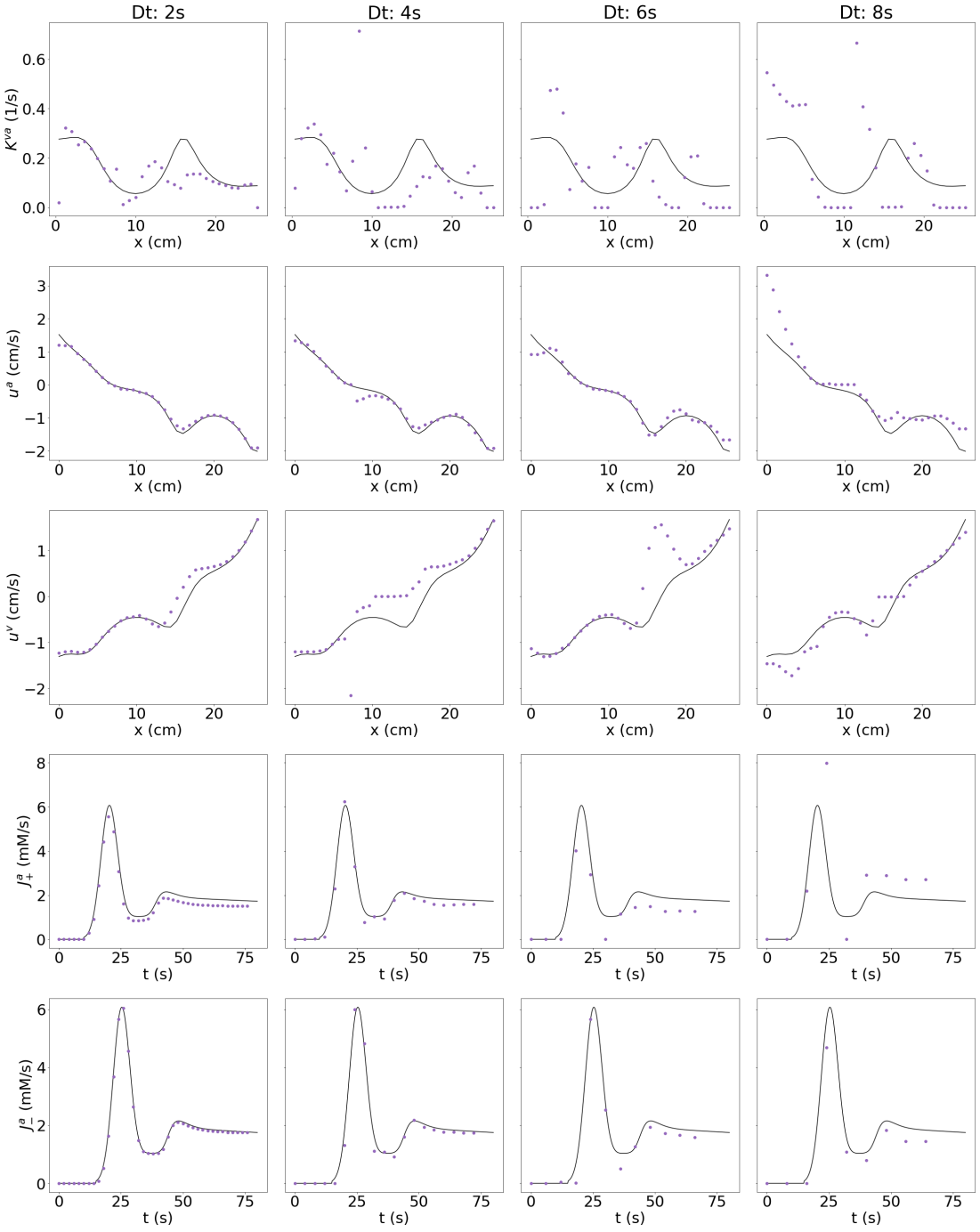

Figure 5.20: Parameter fitting results for time sampling resolution of 2, 4, 6 and 8s in two-compartment Case 3. Here the black lines denote the ground truth parameter values, with the dotted points showing the recovered parameter values.

correlation between parameter groups. Figures 5.24-5.26 indicate some correlation between different parameter types and within parameter types. Correlations can be seen between the boundary flow values with the volume fractions (negative correlation), and boundary concentrations (positive correlation). This is expected due to the determination of flow and velocity up to a scaling constant as discussed (5.2.1), and the system influx definition which uses both boundary flow and concentration.

There are strong diagonal elements in Cases 1 and 3 (Figures 5.24 & 5.26, respectively) between the spatial parameters, indicating a relationship between parameters within voxels and the model output. Within Case 1 (Figure 5.24) a strong diagonal of negative correlation coefficients can be seen between v^v and v^a, indicating values within a voxel have an inverse relationship on the solution. Such an inverse relation would indicate that increases in v^v shows similar outcomes to decreases in v^a and vice versa.

Within Case 2 (Figure 5.25) strong correlation coefficients are reported between different time points for the boundary concentrations (C^a_+, C^a_-). Interpretation of this suggests that increases/decreases of boundary concentrations at earlier and later time points give similar changes to the solution. This could be due to an overall effect on the system of more or less concentration entering over the entire time course. Also indicated in Figure 5.25 are correlations between v^a and v^v values both within and between these parameter groups. These correlations have a strongly symmetric shape indicating correlation within the left and right sides of the system separated by the flow crossing point, suggesting this might be related to the extremely symmetric nature of Case 2. Cases 1 and 3 (Figures 5.24 & 5.25) do not report strong correlations between the boundary concentrations and the spatial parameters as reported for Case 2, potentially due to higher flow and perfusion values present in Case 2.

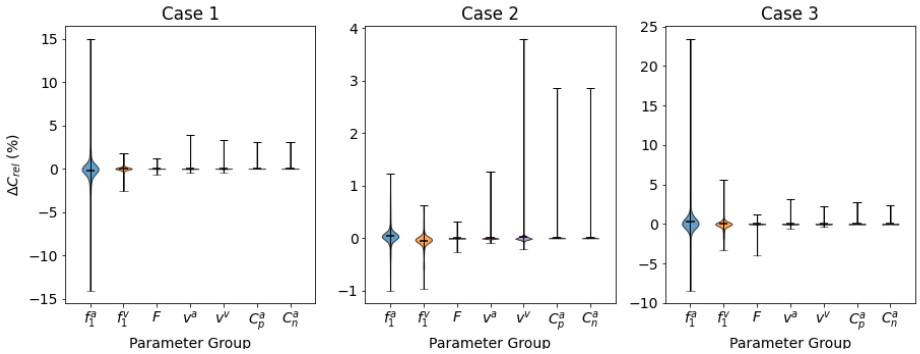

Figure 5.21: Violin plots of resulting changes to the system tissue concentration of adjusting voxelwise parameter values by + 5%. The distributions are shown for each parameter type showing the difference at every coordinate point when adjusting the parameter value.

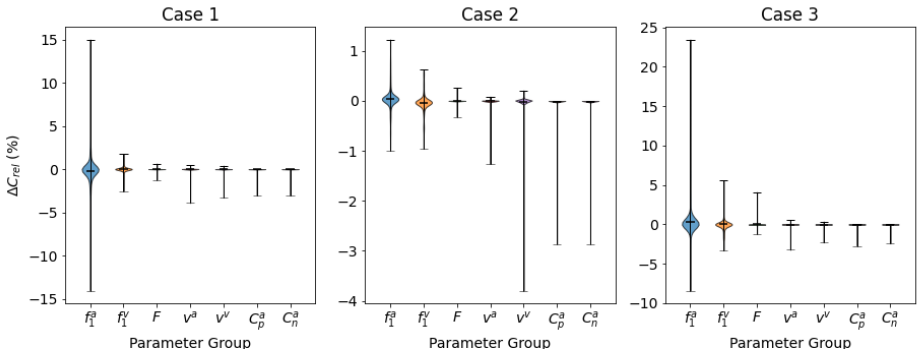

Figure 5.22: Violin plots of resulting changes to the system tissue concentration of adjusting voxelwise parameter values by − 5%. The distributions are shown for each parameter type showing the difference at every coordinate point when adjusting the parameter value.

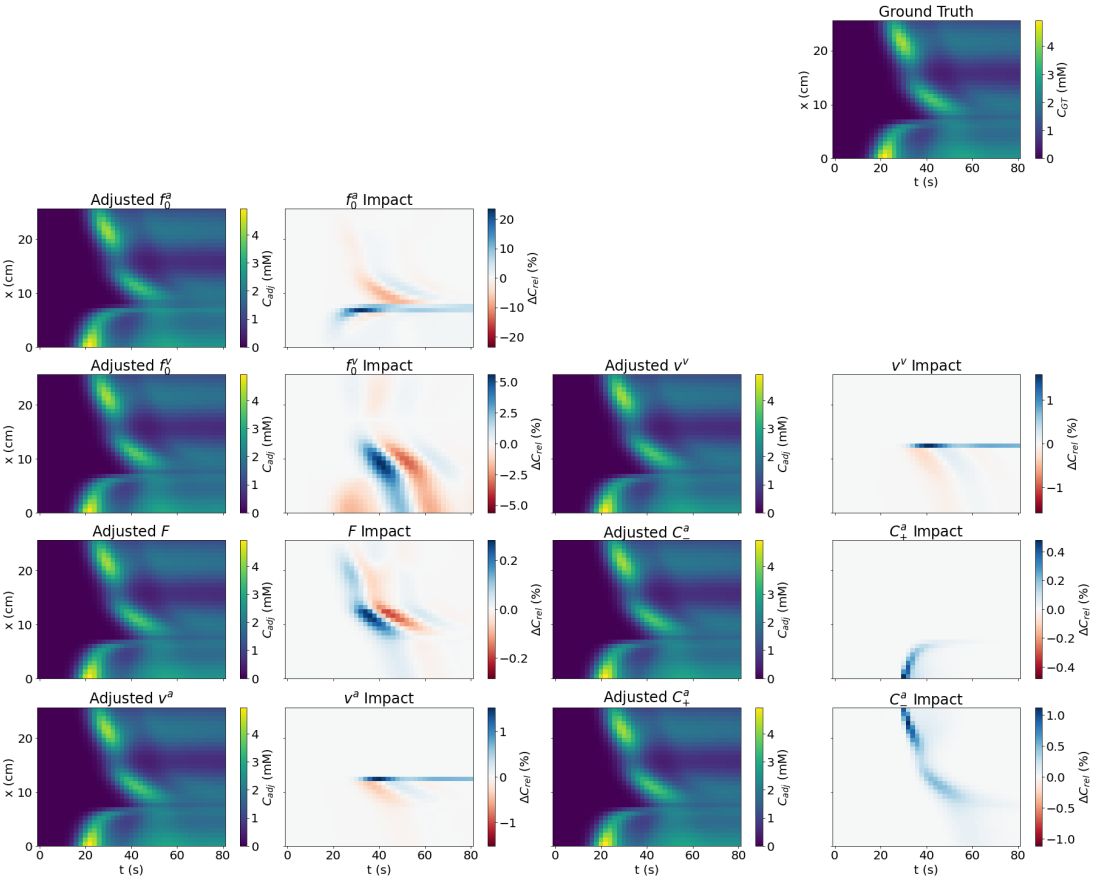

Figure 5.23: Changes in the resulting tissue concentration for the two-compartment sensitivity analysis of Case 3. Changes are coordinate wise one parameter value is adjusted by +5% (x=12.4cm for spatial parameters or t=30s for boundary concentrations) at a time, each pair of panels relates to a different parameter type. Panels show the concentration evolution with the updated parameter (1$^{\text{st}}$ and 3$^{\text{rd}}$ columns) and the relative different as a percentage of the maximal system concentration (2$^{\text{nd}}$ and 4$^{\text{th}}$ columns). The ground truth concentration is shown in the top right hand panel.

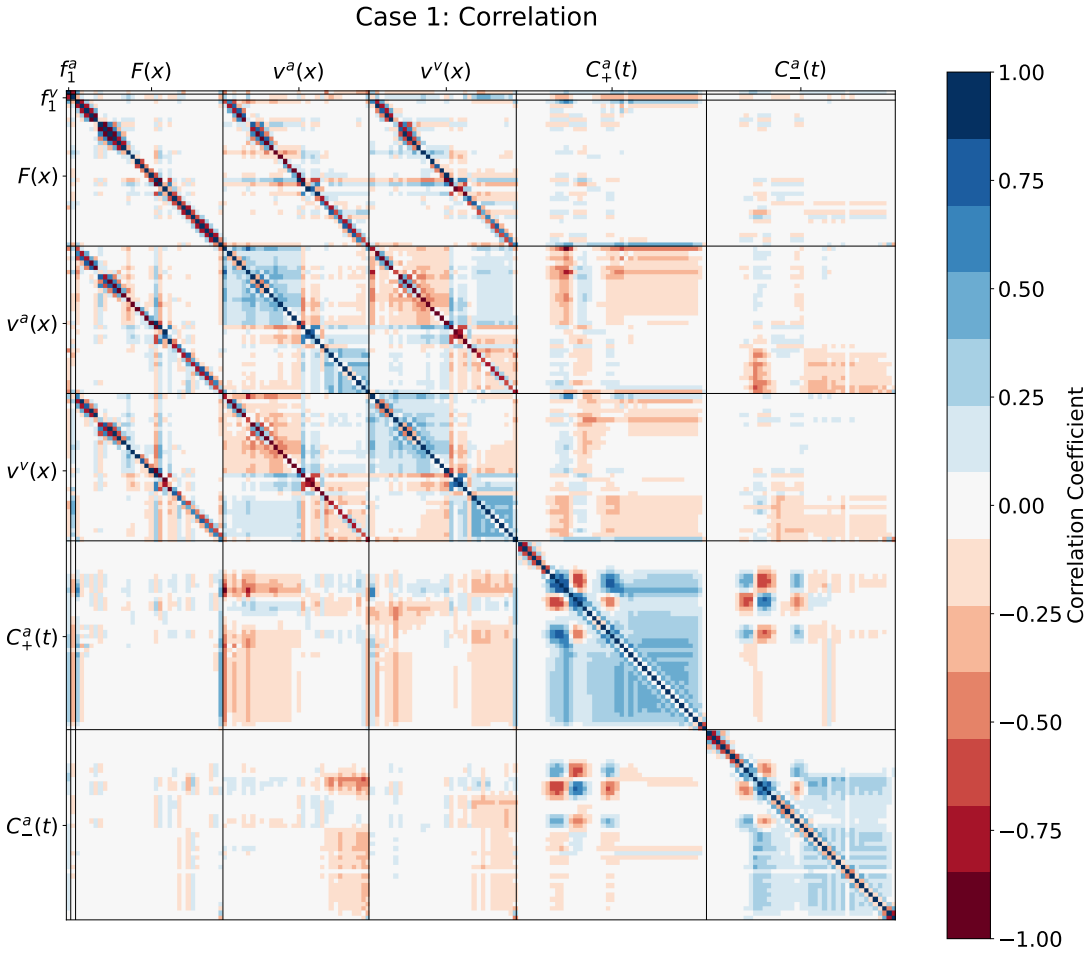

Figure 5.24: Correlation coefficient values derived for model parameters in two-compartment Case 1. Each parameter type is separated by the vertical and horizontal black lines, with every pixel within the plot corresponding to a single parameter e.g. f_0^a or $F(x_i)$.

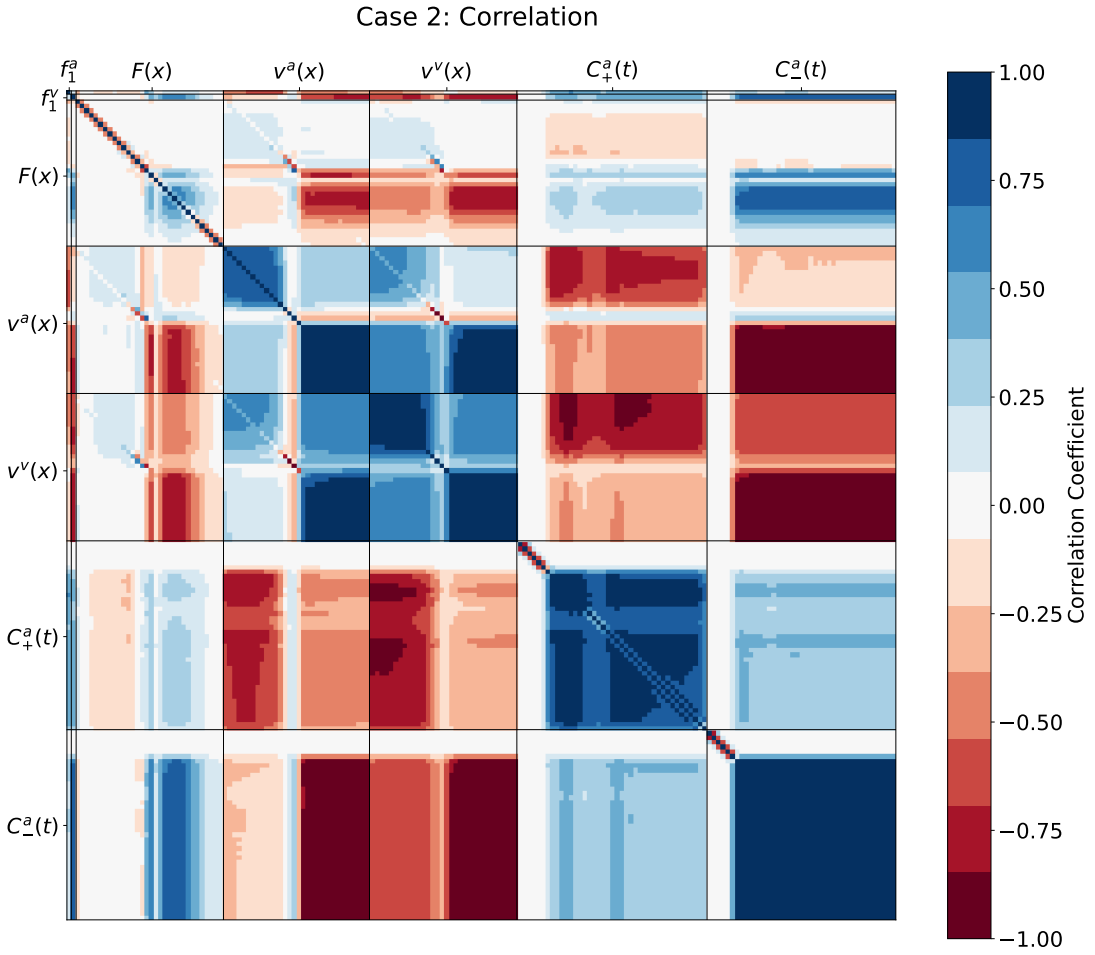

Figure 5.25: Correlation coefficient values derived for model parameters in two-compartment Case 2. Each parameter type is separated by the vertical and horizontal black lines, with every pixel within the plot corresponding to a single parameter e.g. f_0^a or $F(x_i)$.

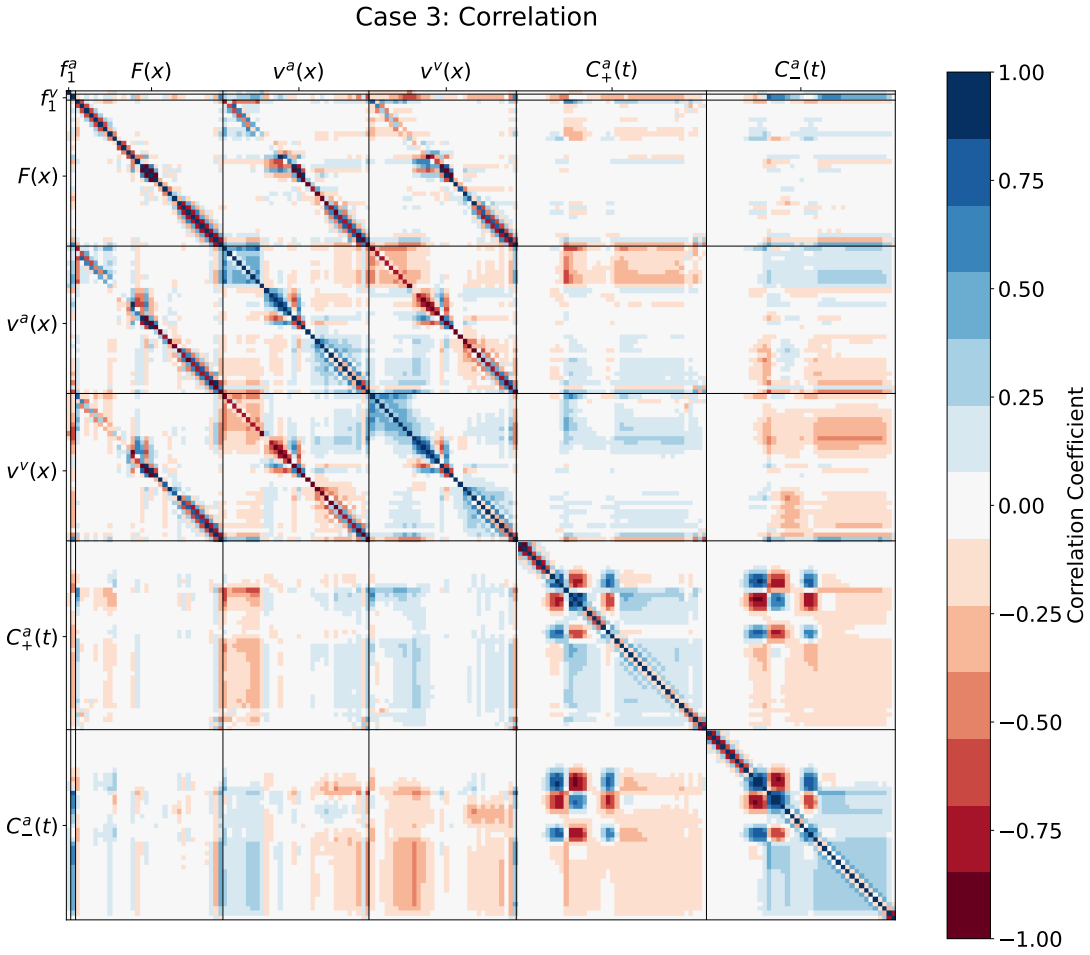

Figure 5.26: Correlation coefficient values derived for model parameters in two-compartment Case 3. Each parameter type is separated by the vertical and horizontal black lines, with every pixel within the plot corresponding to a single parameter e.g. f_0^a or $F(x_i)$.

5.4 Discussion

The aim of this in-silico study was to determine if unconstrained spatio-temporal models for DCE-MRI are fundamentally identifiable. The data indicate that this is the case for one-compartmental blood flow models, but not for two-compartmental perfusion models.

In the absence of significant measurement error, parameters of the one-compartment model can be reconstructed accurately without imposing additional constraints on the model. They are identical even with widely different choices of the initial guesses, suggesting the solution is also unique. Reconstructions of the influxes at the boundary of the system are also accurate, confirming the idea that spatio-temporal models remove the need for a separate measurement of an AIF.

While the analysis in this study used non-linear optimisation, the uniqueness of the one-compartmental solutions aligns with the fact that the model equations can be recast in a linear form (Sourbron, 2014) - in a similar way as for standard temporal one-compartmental models (Flouri *et al.*, 2016). Here implementation of the model in linear form was not chosen as this is known to be more noise-sensitive, and does not translate as easily to the two-compartmental scenario. Measurement error (noise and undersampling) naturally reduces the accuracy and precision of the parameters, but in a predictable and expected manner. One open question is how the parameter accuracy and precision compares to a conventional analysis using measured input functions and standard voxel-by-voxel temporal models.

Parameter reconstructions are significantly less accurate for the two-compartmental perfusion model. Multiple solutions have been found that are compatible with the data, and therefore a single optimal solution cannot be identified using a goodness-of-fit criterion alone. Correlation coefficients also suggest the existence of some strong correlations within and between the different parameter groups effect on the resulting model solutions, particularly within the same voxel. Coupled with the sensitivity analysis which shows that most individual parameters have a minimal impact on solutions, these

interactions may contribute to the reported degeneracy.

One approach to resolving the degeneracy in the two-compartmental model may involve modifying the experimental conditions to increase the structure in the data. However, the options in DCE-MRI are limited. The smallest sampling interval considered in this study was 2 seconds, and therefore there may be some room for improvement by considering faster scan sequences. Beyond that, the only additional variable that can be modified substantially is the injection protocol. The current setup uses a single bolus injection, and while this can easily be modified in practice to split the dose over two injections (Ingrisch & Sourbron, 2013), there is currently little evidence that this translates to more accurate solutions.

Hence this may indicate that additional constraints are needed to pin down multi-compartmental spatio-temporal models. Possible solutions previously proposed for one-compartment systems may well translate to two-compartment systems, such as the use of Darcy flows or other physical constraints to reduce the number of free variables (Naevdal *et al.*, 2016), adding regularisation to impose smoothness of the solution (Zhou *et al.*, 2021; Liu *et al.*, 2021a), fixing less critical parameters to literature values (Pellerin *et al.*, 2007), or reverting to a measured AIF at the boundaries of the imaging slab. The use of physical constraints derived from principles of fluid dynamics and porous media theory presents a particularly attractive approach as it also provides a mechanism for studying the mechanical properties of physiological flow. While such constraints may not be necessary for one-compartmental systems, they may prove essential in the multi-compartment case.

Beyond modifying experimental conditions or imposing additional constraints, another strategy for reducing the degeneracy in the solutions may well be to improve the optimisation itself. Setting suitable initial conditions, for instance, may well help to bias the solution towards the correct value, and may be feasible without loosing generality. For instance, exploratory simulations with the 1D toy models suggest that initial values where the arterial velocity is higher than the venous velocity leads to better parameter recovery than randomly chosen initial values, and this is consistent with physical reality. An

alternative approach, common for instance in other inverse problems in imaging such as coregistration (Studholme *et al.*, 1996; Maes *et al.*, 1999), may be to employ a multi-resolution approach, fitting parameters initially at coarse resolution and then stepwise refining the estimates until the image resolution is reached. Additionally, considering the observation that estimates are most accurate in the arterial parameters, an improvement may be possible by reparameterising the model in terms of the arterial flow field $f^a(x)$ rather than using the perfusion field $F(x)$ as a primary variable. Finally, solutions proposed for temporal model fitting in DCE-MRI may well help in spatiotemporal modelling as well, such as the use of model selection, which can potentially be generalised to a voxel-by-voxel approach, and/or using the results of one-compartment fits to initialise a two-compartment analysis.

The optimal strategy may also depend on the parameter that is the primary interest of the measurement. As shown, results are considerably more reliable in the properties of the upstream (arterial) compartment compared to the distal (venous) compartment and particularly the exchange parameter (perfusion) itself. Hence in clinical applications where the primary aim is to characterise the arterial system, issues of uniqueness identified in this study may be less critical. Unfortunately, the interest in many key clinical applications of perfusion imaging, such as acute stroke (Demeestere *et al.*, 2020) or cancer (van Dijken *et al.*, 2019), is primarily in a measurement of perfusion as this is a key metric to understanding tissue viability or metabolic activity. While the result in this study has shown that measurement of perfusion comes with significant numerical error, the use of spatiotemporal modelling does remove the equally substantial error that comes from assuming a single upstream feeding artery (Calamante *et al.*, 2006). It is currently unknown whether, and to what extent, this offsets the numerical reconstruction errors observed in the spatiotemporal model.

This study is obviously limited by the use a of a one-dimensional toy model. In reality, fully unconstrained spatiotemporal modelling for DCE is only relevant when applied to 3D data, as through-plane exchange of indicator cannot be excluded in realistic scenarios. However, application of

multi-compartmental modelling in 3D comes with significant computational challenges that are currently largely unresolved. Standard gradient-descent type optimisation as performed in this study is unlikely to be practically feasible in 3D, though this has not yet been fully explored. The use of 1D models allows for a flexible exploration of fundamental issues of parameter identifiability, but there is no guarantee that the findings translate to the 3D scenario. Indeed, 3D data are significantly more entangled due to the spatial connections in the other dimensions, and this may well help to resolved any degeneracies found in 1D. Future studies should therefore focus in the first place on developing computational methods that are able to solve spatiotemporal two-compartment models in reasonable computation times, before the issue of parameter identifiability can be investigated in-silico in 3D data. Recent developments in deep learning, specifically the use of PINNs and their successful application in related problems, has offered some hope that a solution may be technically feasible.

5.5 Conclusions

This chapter provides proof of concept that one-compartmental blood flow models are fully identifiable and do not require a separate measurement of the AIF. Arterial properties of two-compartmental perfusion models have comparable accuracy but perfusion fields and venous flows are cannot be measured reliably. Future studies should focus on exploring the use of physical constraints, improved optimisation and on development of computational solutions for the 3D case.

Chapter 6

Physics Informed Neural Networks for Tracer Kinetics

6.1 Introduction

Throughout the research presented in this thesis, the sheer number of parameters needed to characterise a spatiotemporal TKs model present a constant challenge. A step change in the number of parameters is needed to increase the compartmental complexity, and is also a prerequisite for moving to higher dimensional systems (Section 4.6 and Section 5.4). The advancement of methodologies that can handle this parameter increase is essential for widespread application of complex spatiotemporal models to 3D clinical perfusion MRI data.

Within Chapter 5, the classical gradient descent technique for the inverse approach shows a good reconstruction of system influxes and compartmental velocities for noiseless and high SNR measurements (Section 5.3.2). However, the runtime even at this 1D proof of concept stage would become quickly prohibitive when moving towards 2D and 3D systems (Section 5.4). There is a motivation for moving directly to 3D systems rather than bridging this gap via 2D systems. When addressing clinical DCE-MRI scan data, erroneous spatiotemporal parameter fitting will occur if only a 2D slice is considered due to the unaccounted indicator transport into or out of the slice. As a consequence, the spatiotemporal problem requires a new optimisation approach designed for large numbers of parameters and large data sets.

More recently, machine learning has been applied in the field of spatiotemporal

TKs with a convolutional NN approach proposed by Liu *et al.* (2021b). This method utilises a U-Net (Ronneberger *et al.*, 2015) based architecture where the parameters are evolved through a finite volume simulation framework and compared to the DSC-MRI data. The approach by Liu *et al.* (2021b) is aimed at one-compartment systems with both convective and diffusive transport. Although this is not a PINNs implementation, the robustness of the reported parameters in the presence of noise certainly adds promise to an approach based on neural networks.

The PINNs framework was proposed by Raissi *et al.* (2017a,b). This was paradigm-shifting work which the authors proceeded to develop: providing solutions to the Navier-Stokes equations specifically (Raissi & Karniadakis, 2018) and later presenting a generalised framework exploiting the methodology for both forwards and inverse problems (Raissi *et al.*, 2019). The PINNs use residual functions based on the physical laws of the system, for example, mass conservation or incompressibility of flow. Recent developments in the field of PINNs include: adaptive activation functions for use in both forward and inverse modelling (Jagtap *et al.*, 2020); and the use of a deep neural network to interpret spatio-temporal data (Koeppe *et al.*, 2020).

The subsequent use of PINNs has been widespread and numerous (Sharma *et al.*, 2023). Raissi *et al.* (2018, 2020) further showcase the capability of PINNs through the sole use of concentration data to solve for 'hidden' or unmeasurable velocity and pressure fields governing the Navier-Stokes equations. This latter study (Raissi *et al.*, 2020) is the most promising approach in the application of PINNs for the recovery of spatiotemporal TK parameters. Other studies have established the suitability of this approach, with real clinical data applied to extract arterial wall pressures from 4D flow MRI via a PINNs-based method (Kissas *et al.*, 2020). Kissas *et al.* (2020) uses sparse measurement of blood velocity and arterial vessel cross sectional area within a simplified pulsatile flow model. More recently, van Herten *et al.* (2022) have applied PINNs for standard isolated voxel fitting of DCE-MRI data. This study (van Herten *et al.*, 2022) applied the 2CXM to myocardial blood flow DCE-MRI data.

To the best of the author's knowledge, application of PINNs to a spatial

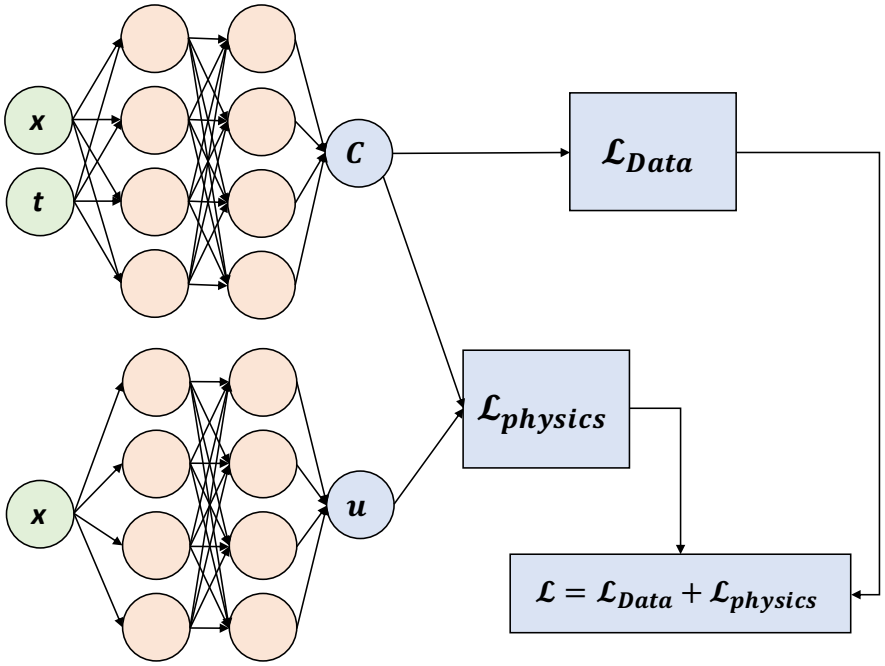

Figure 6.1: Illustrative network setup for convection only one-compartment systems. The number of internal nodes and hidden layers are reduced for clarity.

multi-compartment modelling setup has yet to be realised. Usually, approaches applying PINNs for fluid mechanics or MRI-imaging purposes consider systems comparable to a one-compartment setup. In these cases, the concentration or similar transport marker is directly accessible by measurement (Kissas *et al.*, 2020) or voxels are assessed as isolated systems (van Herten *et al.*, 2022).

The overall aims of this chapter are three-fold. Firstly, to produce a PINN implementation to recover spatiotemporal TK parameters for both one and two-compartment systems in 1D. Secondly, to investigate the robustness of the parameter recovery outcomes in the presence of increasing measurement noise and sampling intervals. Finally, a one-compartment 2D system parameter is also detailed with limited methodology adjustment to illustrate the straightforward extension of the developed PINN implementation.

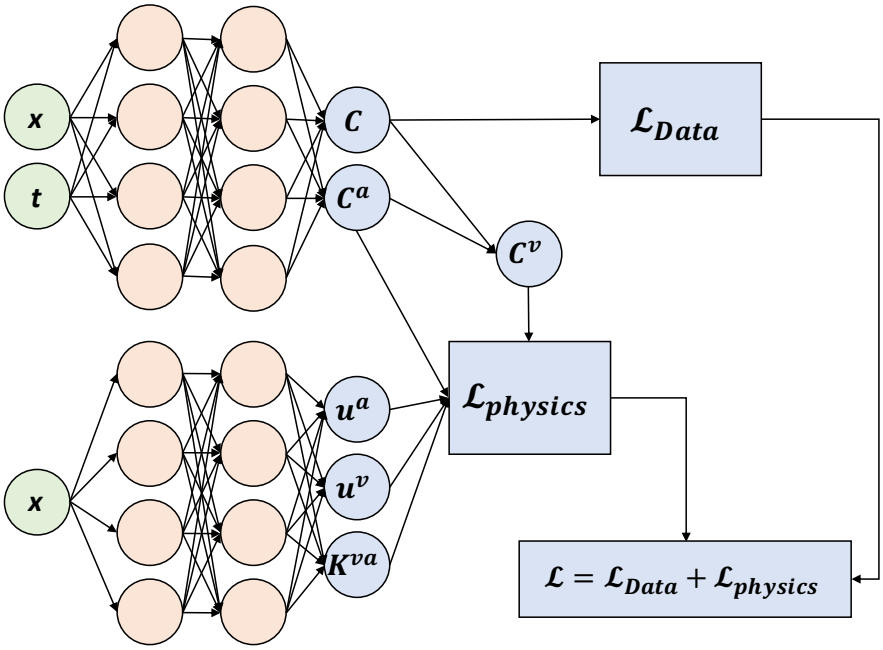

Figure 6.2: Illustrative network setup for convection only two-compartment systems. The number of internal nodes and hidden layers are reduced for clarity.

6.2 Methodology

6.2.1 Network Architecture

The network architecture follows similar setup to the proposed network detailed by Raissi *et al.* (2020). The implementation of this work is freely available via GitHub under the MIT license[1]. These implementations have been restructured for compatibility with Tensorflow version 2 (Abadi *et al.*, 2015) by the Centre for Environmental Modelling And Computation (CEMAC) at the University of Leeds. These updated implementations of the original work by Raissi *et al.* (2019, 2020) are presented as a set of jupyter iPython notebooks hosted on the CEMAC GitHub[2] under the Creative Commons Attribution 4.0 International license.

The setup to apply PINNs with TK equations uses a modified version of the CEMAC notebook produced to implement Raissi *et al.* (2020). Most notably, the spatiotemporal TK PINN (tkPINN) implementation uses two separate neural networks that are optimised in tandem via the same loss function. The networks are separated due to the vastly different parameter field sizes. The first network

[1] https://github.com/maziarraissi/HFM/tree/master
[2] https://github.com/cemac/LIFD_Physics_Informed_Neural_Networks/tree/main

takes both spatial and temporal inputs and so returns spatiotemporal fields as an output. The second takes only a spatial input and as such returns outputs that are constant in time. An overview of this general setup is shown for the one- and two-compartment system parameters in Figures 6.1 and 6.2.

6.2.2 Loss Terms

Within the tkPINNs method, there are two passes through the network to produce the data and physics loss terms for the overall loss function. First, the data coordinate points or "experimental points" are passed through the networks and the resulting tissue concentration is compared to the measurement tissue concentration. Secondly, the physics coordinate points or "collocation points" are passed through the networks with all resulting parameters used to define several residual equations defining the 'hidden' TKs of the system. These points are denoted as collocation points as numerical solutions of differential equations are known as collocation methods. It is not necessary to have matching experimental and collocation points. In instances with sparse data it is usual to have increased resolution for the physics coordinates (Raissi *et al.*, 2020).

For the one-compartment systems, physics residual equations for the tkPINNs setup are as follows:

$$\Phi_1 : 0 = \frac{\partial C}{\partial t} + \vec{\nabla} \cdot (uC) \tag{6.1}$$

$$\Phi_2 : 0 = |C| - C \tag{6.2}$$

$$\Phi_3 : 0 = 1 - C - d \tag{6.3}$$

Here, d is defined by $d = 1 - C$ to be the adjoint of the tissue concentration. This extra parameter is applied in Raissi *et al.* (2020) to help guide the network optimisation. With the residuals similarly given for the two-compartment systems

as:

$$\Phi_1 : 0 = \frac{\partial C^a}{\partial t} + \vec{\nabla} \cdot (u^a C^a) + K^{va} C^a \tag{6.4}$$

$$\Phi_2 : 0 = \frac{\partial C^v}{\partial t} + \vec{\nabla} \cdot (u^v C^v) - K^{va} C^a \tag{6.5}$$

$$\Phi_3 : 0 = C - C^a - C^v \tag{6.6}$$

$$\Phi_4 : 0 = |C^a| - C^a \tag{6.7}$$

$$\Phi_5 : 0 = |C^v| - C^v \tag{6.8}$$

$$\Phi_6 : 0 = |K^{va}| - K^{va} \tag{6.9}$$

$$\Phi_7 : 0 = 1 - C - d \tag{6.10}$$

Within the residual terms any differential terms are calculated using the inbuilt automatic differentiation functions in Tensorflow (e.g. $\frac{\partial C}{\partial x}$=`tf.gradient(C,x)`). These residuals are collected into the physics loss term using the mean squared error of each residual:

$$\mathcal{L}_{Physics} = \sum_i \mu(\Phi_i^2) \tag{6.11}$$

Similarly, the data loss term is calculated using the mean squared error between the network output concentration, C^{Net}, and the measurement concentration, C:

$$\mathcal{L}_{Data} = \mu((C^{Net} - C)^2) \tag{6.12}$$

Within a simple PINNs setup there is enforcement of soft constraints only. Soft constraints are allowed to be violated as they are merely a part of the loss function to be minimised rather than an enforced hard constraint. For flow incompressibility the soft constraint would prescribe:

$$min(\nabla \cdot \vec{f^a}) \tag{6.13}$$

However, this minimisation (Equation 6.13) cannot be guaranteed to equal exactly zero - as would be the case for a true incompressible system. To avoid problems stemming from the soft constraint, fitting is carried out in the tissue

concentration parameters (u, u^a, u^v, K^{va}).

For network training both spatial and temporal inputs are supplied in dimensionless form. The input array sizes must match so for a given spatial input length N and temporal input length T a mesh grid is created (NxT) which is then flattened giving two (NTx1) inputs. Collocation and experimental points are produced separately for the physics and data loss. For investigation of fully sampled measurements (Dt=2s) the same set of coordinate points is used to evaluate the data and physics loss terms. For undersampled measurements (Dt>2s), the experimental points for the data loss term are reduced accordingly, while the physics collocation points remain at 2s intervals. The ordering of the data and physics training points are randomised for each separate iteration.

The only data supplied to the network is for the tissue concentration. Therefore, the only way the constrain the arterial and venous concentration is through the physics residual equations that contribute to the loss function.

The network detailed can be easily adjusted to handle problems with additional dimensions. To demonstrate this the tkPINNs is extended to include a y-coordinate. The physics residuals can be extended to reflect advective transport in both x and y-directions.

6.2.3 Models for Investigation

The outcome of the PINN methodology is investigated on systems setup in both one- and two-dimensions.

One-Dimension Models

The tkPINNs method is investigated on 1D for a one-compartment convective blood flow model, and a two-compartment perfusion model with convective blood transport. Definitions of both the one-compartment and two-compartment models are set out in Chapter 5 (Section 5.2) based on the theoretical framework as defined in Chapter 2 (Section 2.4). The ground truth data is produced via the same finite difference scheme using the local concentration picture defined from Equations 5.12-5.16. The same 3 systems of

each model type defined in Chapter 5 are run using the tkPINNs method, to enable comparisons to be made.

Two-Dimension Models

The tkPINNs method is further investigated in 2D for a one-compartment convective blood flow model. The theoretical definitions match to general equations used for the one-compartment convective blood flow model in Chapter 5.

After discretisation, the one-compartment spatiotemporal model in two dimensions reduces to an $N_x N_y$-compartment temporal model (Sourbron, 2014):

$$\begin{aligned} c_{i,j}(t+\Delta t) =& c_{i,j}(t) + \frac{\Delta t}{v_{i,j}}((k_{ij,i-1j}c_{i-1,j}(t) + k_{ij,i+1j}c_{i+1,j}(t) \\ &+ k_{ij,ij-1}c_{i,j-1}(t) + k_{ij,ij+1}c_{i,j+1}(t) \\ &- (k_{i-1j,ij} + k_{ij-1,ij} + k_{i+1j,ij} + k_{ij+1,ij})c_{i,j}(t)) \end{aligned} \tag{6.14}$$

$$k_{ij,i-1j} = \begin{cases} \frac{f_{x,ij}}{\Delta x}, & f_{x,ij} > 0 \\ 0, & f_{x,ij} \leq 0 \end{cases} \qquad k_{i-1j,ij} = \begin{cases} 0, & f_{x,ij} \geq 0 \\ \frac{-f_{x,ij}}{\Delta x}, & f_{x,ij} < 0 \end{cases} \tag{6.15}$$

$$k_{ij,ij-1} = \begin{cases} \frac{f_{y,ij}}{\Delta y}, & f_{y,ij} > 0 \\ 0, & f_{y,ij} \leq 0 \end{cases} \qquad k_{ij-1,ij} = \begin{cases} 0, & f_{y,ij} \geq 0 \\ \frac{-f_{y,ij}}{\Delta y}, & f_{y,ij} < 0 \end{cases} \tag{6.16}$$

Here the flow in the x-direction $f_{x,ij}$ is defined at the left interface of voxel ij, and flow in the y-direction $f_{y,ij}$ is defined at the bottom interface of voxel ij. Total outgoing transport from a voxel is $k_{ij} = k_{i-1j,ij} + k_{i+1j,ij} + k_{ij-1,ij} + k_{ij+1,ij}$. Additional free parameters of the model are the concentrations $c_l(y,t)$, $c_r(y,t)$, $c_t(x,t)$ and $c_b(x,t)$ at the left, right, top and bottom boundary of the system, respectively. For numerical stability, the time step Δt must be chosen to be

Table 6.1: Ground truth values for the 2D one-compartment system. All x values used are in cm. P_{AIF} is a population AIF (Parker *et al.*, 2006) with a defined delay (d) and a scaling factor ($0 \leq s_f \leq 1$). $G(c, h)$ denotes a Gaussian with center (c) and height (h).

Parameters	Case 1
$v(ml/ml)$	$0.3sin^2(0.2x) + 0.3cos^2(0.2y)) + 0.3$
$f_{x,0j}(ml/s/cm^2)$	1.2
$f_y(ml/s/cm^2)$	$0.3sin^2(0.2x) + 0.2cos^2(0.15y) + 0.25$
$c_l(mM)$	$P_{AIF}(8s, G(0.5L_y, 1))$
$c_b(mM)$	$P_{AIF}(5s, G(0.5L_x, 1))$
$c_r(mM)$	0
$c_t(mM)$	0

smaller than the smallest voxel mean transit time:

$$\Delta t < \min_i \left\{ \frac{v_{ij}}{k_{ij}} \right\} \tag{6.17}$$

In a 2D scenario, the incompressibility of flow restricts the flow pattern. An incompressible flow is produced via the definition of the left-hand boundary x component of the flow, with the y component defined at all voxels. The remaining x component flow values are determined by obeying incompressibility as:

$$k_{ij,ij-1} + k_{ij,i-1j} + k_{ij,ij+1} + k_{ij,i+1j} = k_{ij-1,ij} + k_{i-1j,ij} + k_{ij+1,ij} + k_{i+1j,ij} \tag{6.18}$$

The 2D system ground truth values are shown in Table 6.1.

6.2.4 Setting hyperparameters

As with many optimisation strategies, several hyperparameters must be chosen through the tkPINNs method. For the tkPINNs method presented the hyperparameters used are defined in Table 6.2. A full investigation of the interplay and optimisation of hyperparameter choice is out of the scope of this thesis.

The initial weights in the network are assigned drawing from a normal Gaussian distribution, implemented using the numpy `random.normal` module. For reproducibility the random seed is assigned to ensure repeat runs are identical for the same computational setup.

Table 6.2: Hyperparameters used within the tkPINNs setup across different system types.

	System Type		
Hyperparameters	One-compartment 1D	Two-compartment 1D	One-compartment 2D
Hidden Layers	10		
Nodes per Hidden Layer	100	150	100
Learning Rate	1×10^{-3}		
Training Iterations	7,500	250,000	100,000
Batch Size	NxT		
Activation Function	Swish		
Optimisation Strategy	Tensorflow Adams Optimiser		
Initial Weights	Random allocation for normal distribution		
Loss Function Weighting	Unity		

6.2.5 Simulation Systems

The tkPINNs setup defined in this chapter is applied to the 3 example cases for both the one- and two-compartment models previously investigated in Chapter 5. The 2D one-compartment system is defined based on parameters in Table 6.1, with system dimensions of 25.6×25.6cm over 80s for $\Delta x = \Delta y = 0.8$cm. For each 1D case the PINNs method reports the parameter fitting for noiseless data with 2 second temporal resolutions. Sensitivity to measurement noise was investigated using SNR levels of 5, 10, 15, 20. Sensitivity to measurement undersampling was assessed by applying noiseless data with temporal sampling resolution of 2, 4, 6, 8, 10s. For the single 2D one-compartment blood flow system, parameter fitting is investigated for noiseless data with 2 second temporal resolution.

Each instance of the tkPINNs methodology was run across 16 CPUs (Broadwell E5-2650v4 CPUs 2.2GHz)[1] until the maximum iteration was reached. All main runs were completed using the settings defined in Section 6.2.4.

As in Chapter 5, the $\bar{E}_{rel}$ (Equation 5.19) error metric is applied to determine the accuracy of recovered parameters gained from the tkPINNs method.

6.3 Results

Parameter recovery results from the tkPINNs method in both one-compartment and two-compartment systems across all test cases are summarised in this section and compared to the ground truth parameter values. Additional data tables of $\bar{E}_{rel}$ values are included in Appendix C.

6.3.1 Hyperparameter Tuning

To decide the maximum training iterations needed, several test runs for one and two-compartment 1D cases using a high number of training iterations were carried out. To extract the performance of the network over the initial training iterations both $\bar{E}_{rel}$ values (Figures 6.3 and 6.4), and the total loss function values

[1]Provided by arc3, a research computing resource provided by the University of Leeds.

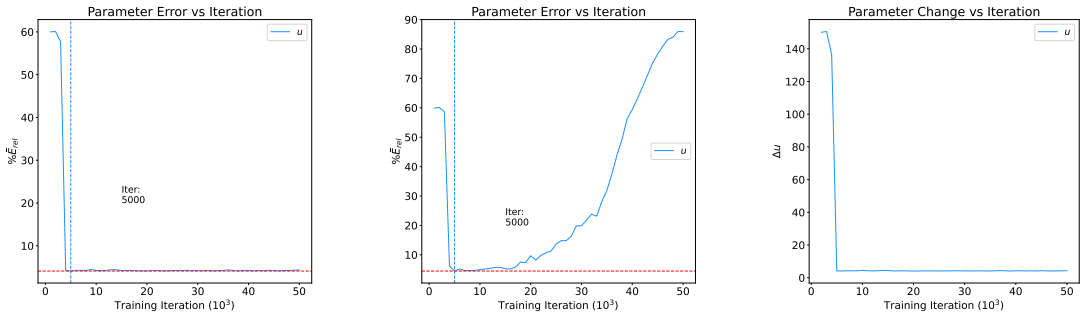

Figure 6.3: $\bar{E}_{rel}$ values across 50,000 test training iterations for one-compartment systems in 1D. Left: Case 1 with noiseless data, middle: Case 1 with measurement SNR 5, and right: parameter change for Case 1 with noiseless data.

(Figures 6.6 and 6.7) were investigated.

For the 1D one-compartment systems, Case 1 is shown using 50,000 test iterations for both a noiseless and noisy system. Figure 6.3 details the $\bar{E}_{rel}$ in both noise regimes, as well as the average update to u at each iteration in the noiseless case. A minimum in $\bar{E}_{rel}$ is observed after 5,000 iterations in both investigations. As parameter updates in the noiseless case also stayed constant after this point, 7,500 iterations were chosen to be comfortably in this optimised region. All further 1D one-compartment system were run to 7,500 training iterations.

For the 1D two-compartment system, both Cases 1 and 3 are shown over 500,000 test iterations for noiseless measurement data. The $\bar{E}_{rel}$ values across u^a, u^v and K^{va} are shown (Figure 6.4) where a trend of improving followed by worsening parameter accuracy is seen most strongly in Case 3. Interestingly, over this area, the loss function shows a decrease (Figure 6.7) while the solution becomes less optimal. To inform an appropriate stopping iteration, the number of iterations needed before a loss function decrease is extracted (Figure 6.5). This shows that the area between 200,000 and 300,000 sees increases above 1,000 iterations before descent. Therefore, training iterations were set to 250,000 for all 1D two-compartment systems.

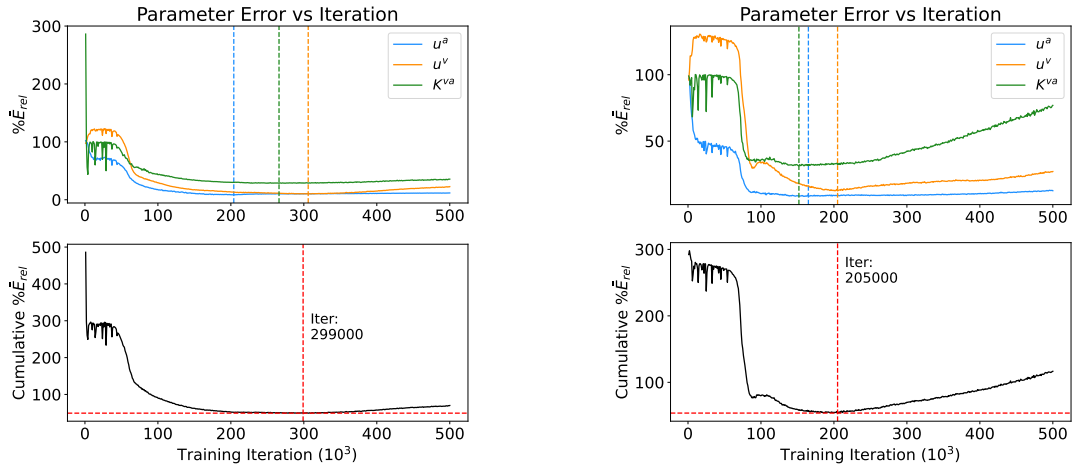

Figure 6.4: $\bar{E}_{rel}$ values of the two-compartment systems across 500,000 test training iterations. left: Case 1 with noiseless data and right: Case 3 with noiseless data.

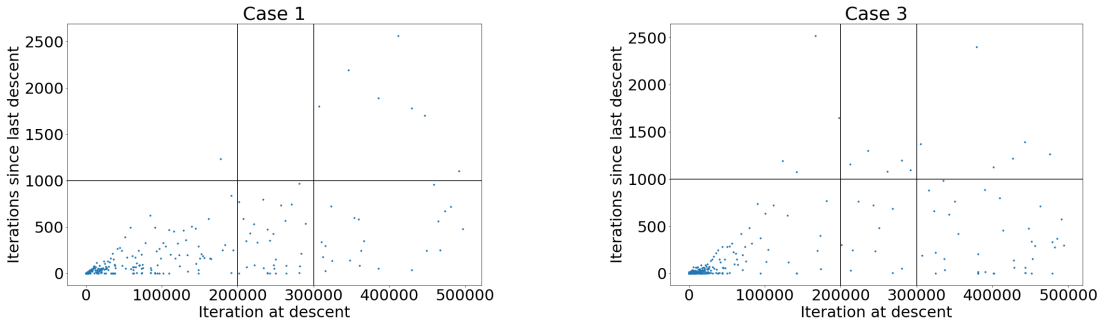

Figure 6.5: Number of iterations since last loss function descent against iteration at these descent points over the 500,000 training iterations in two-compartment systems. Left: Case 1 with noiseless data and right: Case 3 with noiseless data.

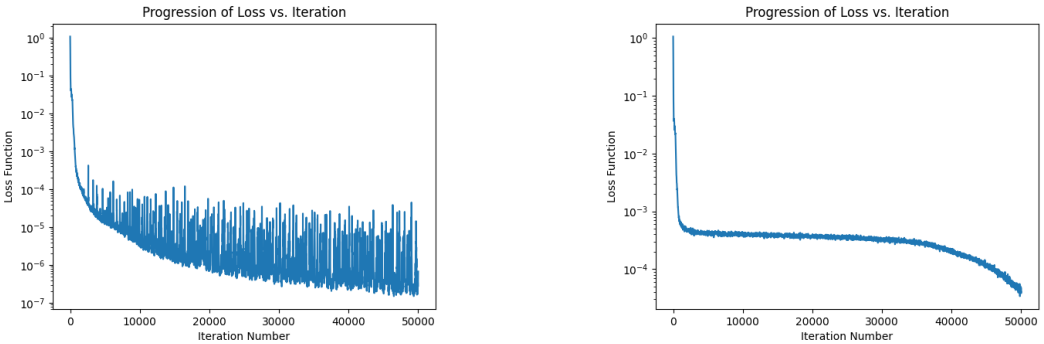

Figure 6.6: Progression of the loss function in one-compartment systems in 1D over the 50,000 training iterations. Left: Case 1 with noiseless data, and right: Case 1 with measurement SNR 10.

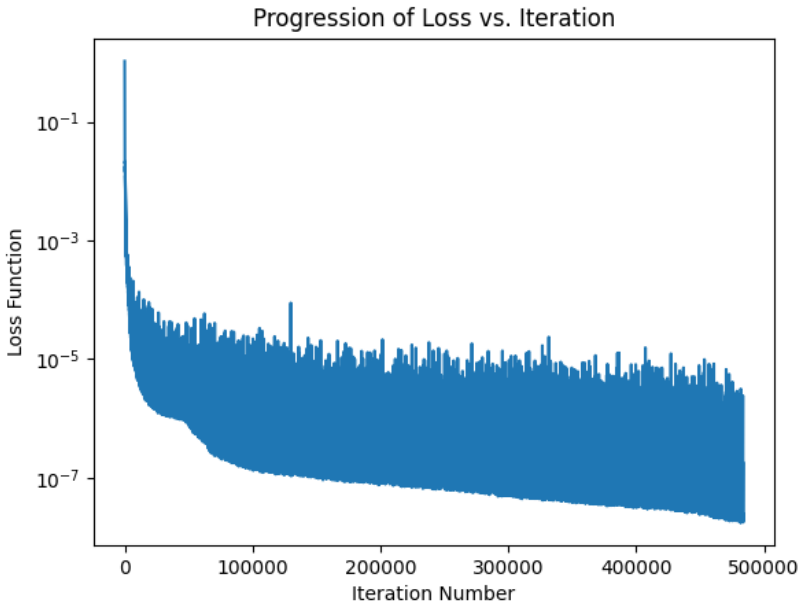

Figure 6.7: Progression of the loss function in two-compartment Case 1 from the tkPINNs network with swish activation over iteration number during training.

6.3.2 One-compartment systems in One Dimension

Results from the one-compartment systems in 1D are detailed here, summarising the influence of system noise and measurement sampling. Runtimes for the one-compartment systems in 1D were between 15 and 35 minutes.

Noiseless Systems

When evaluating noise-free measurement data the tkPINNs method recovers parameters to a low error, reflected in the close match seen to the ground truth parameters (Figure 6.8). This is also reflected in the low $\bar{E}_{rel}$ values seen across all one-compartment cases both for the overall distribution (Figure 6.9) and the voxelwise values (Figure 6.10). The mean $\bar{E}_{rel}$ values are in the range of $4.3 \pm 4.4\%$ to $6.7 \pm 5.5\%$. The variation of $\bar{E}_{rel}$ across the system (Figure 6.10) shows some trend towards higher error in higher velocity places, which follows partially from the definition of $\bar{E}_{rel}$ using the percentage of the parameter mean value rather than the percentage of ground truth at each point. The behaviour is relatively consistent across each case (Figure 6.8).

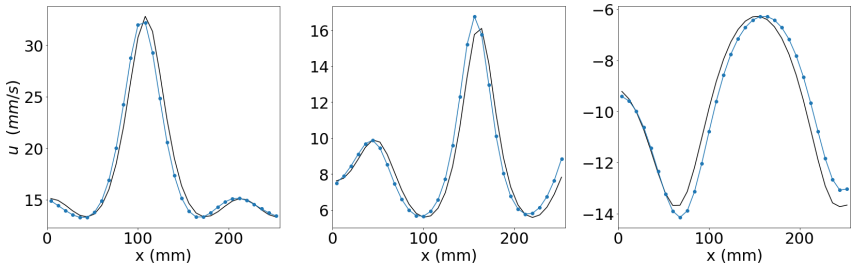

Figure 6.8: Parameter retrieval for one-compartment 1D noiseless measurement data. The black and blue data sets show the ground truth and retrieved parameters, respectively. Cases shown left to right are 1, 2, 3, respectively.

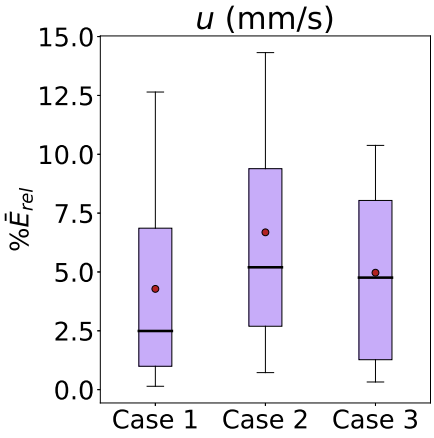

Figure 6.9: Boxplots illustrating the distribution of u $\bar{E}_{rel}$ values for each one-compartment 1D case when applying noiseless measurement data. The whiskers denote the 5^{th}-95^{th} percentile, with the box region indicating the interquartile range. The medians are shown as the solid black line with the means shown as dark red points.

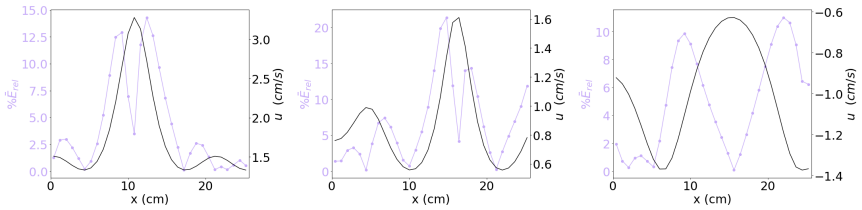

Figure 6.10: Spatial variation in $\bar{E}_{rel}$ values (purple) for each 1D one-compartment case compared to the ground truth values (black) when applying noiseless measurement data. To allow for comparison of different scales, the left hand axis shows the $\bar{E}_{rel}$ range, while the right hand axis shows the ground truth parameter range.

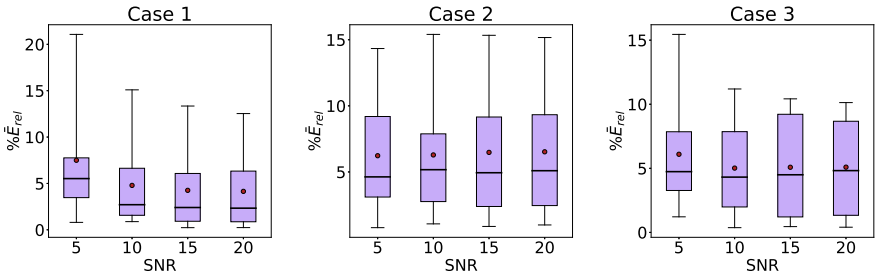

Figure 6.11: Boxplots illustrating the distribution of $\bar{E}_{rel}$ values for each 1D one-compartment case with varying measurement SNR. The whiskers denote the 5^{th}-95^{th} percentile, with the box region indicating the interquartile range. The medians are shown as the solid black line with the means shown as dark red points.

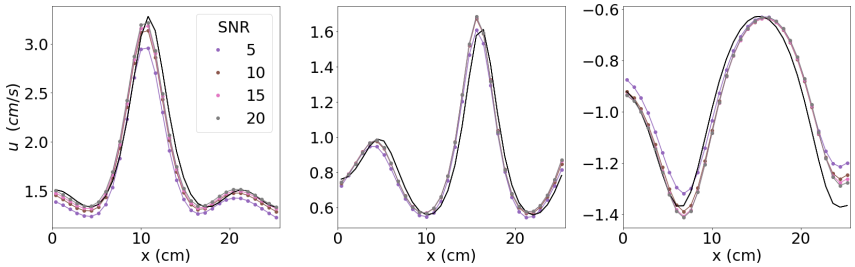

Figure 6.12: 1D one-compartment systems parameter retrieval results from tkPINNs for 7,500 training iterations for measurement data with SNR 5, 10, 15 and 20. The noiseless case parameters are also displayed. Each noise level is displayed in a different colour.

Measurement Noise

The resulting impact on decreasing measurement noise is summarised in Figure 6.11 showing boxplot distributions of the $\bar{E}_{rel}$ values across varying SNR levels. Figure 6.11 shows a robust response to SNR with mean $\bar{E}_{rel}$ values staying largely constant across SNR 5 to 20, showing comparable values to the noiseless recovery. This is reflected by the very similar u retrievals shown against the ground-truth values in Figure 6.12.

Temporal Undersampling

The resulting impact on decreasing measurement sampling is summarised in Figure 6.13 showing boxplot distributions of the $\bar{E}_{rel}$ values across varying sampling intervals. Figure 6.13 shows a robust response to undersampling with mean $\bar{E}_{rel}$ values staying largely constant (5-6%) between sampling intervals of 4 to 8s, showing comparable values to the 2s recovery. This is reflected by the

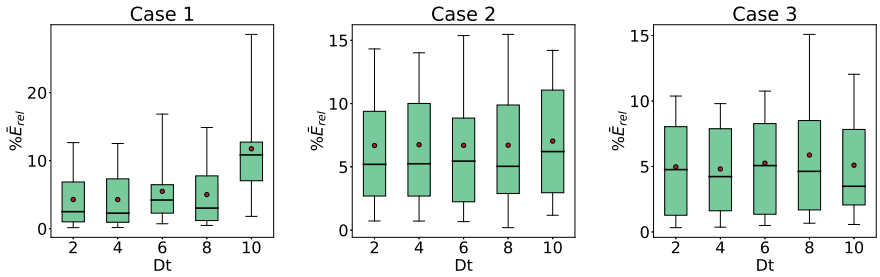

Figure 6.13: Boxplots illustrating the distribution of $\bar{E}_{rel}$ values for each 1D one-compartment case with varying measurement undersampling. The whiskers denote the 5^{th}-95^{th} percentile, with the box region indicating the interquartile range. The medians are shown as the solid black line with the means shown as dark red points.

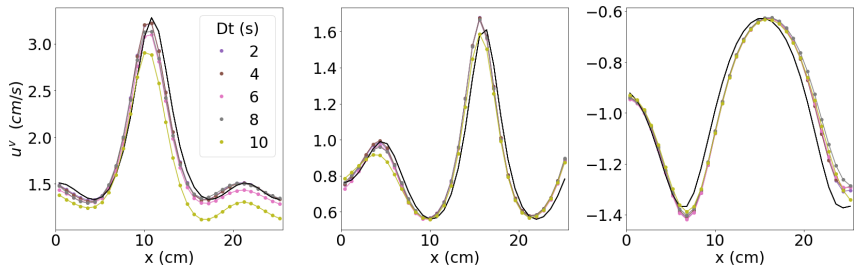

Figure 6.14: 1D one-compartment parameter retrieval results from tkPINNs for 7,500 training iterations for measurement data with temporal sampling measurements of 2, 4, 6, 8 and 10s. The ground-truth case parameters are also displayed. Each sampling level is displayed in a different colour.

very similar u retrievals shown against the ground-truth values in Figure 6.14.

6.3.3 Two-compartment systems in One Dimension

Results from the two-compartment systems are detailed here, summarising the influence of system noise and measurement sampling. Runtimes for the two-compartment systems in 1D were between 6 and 8 hours.

Noiseless Systems

The 3 test cases reported varied parameter retrieval outcomes using the tkPINNs framework when evaluating systems with no measurement noise. Figure 6.15 details the parameter recovery result across each case separately. In Cases 1 and 3, the parameters were recovered close to the ground truth values, while Case 2 showed strong deviations. Distributions of $\bar{E}_{rel}$ are displayed for each test case for u^a, u^v and K^{va} separately, as shown in Figure 6.16. For Cases 1 and

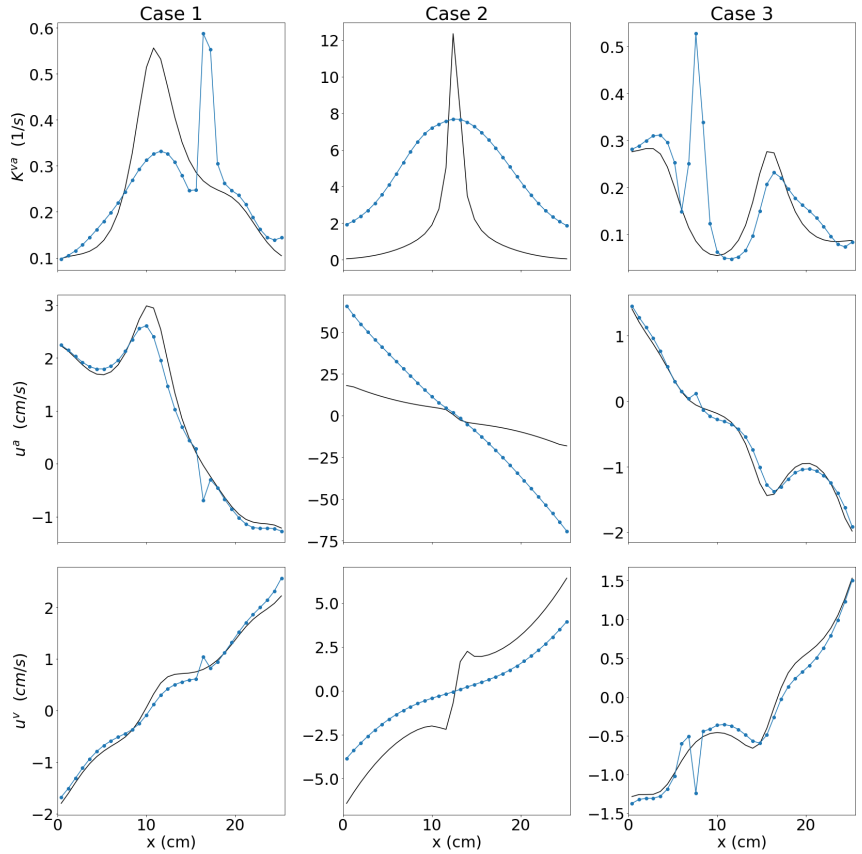

Figure 6.15: Parameter retrieval for 1D two-compartment noiseless measurement data. The black and blue data sets show the ground truth and retrieved parameters, respectively.

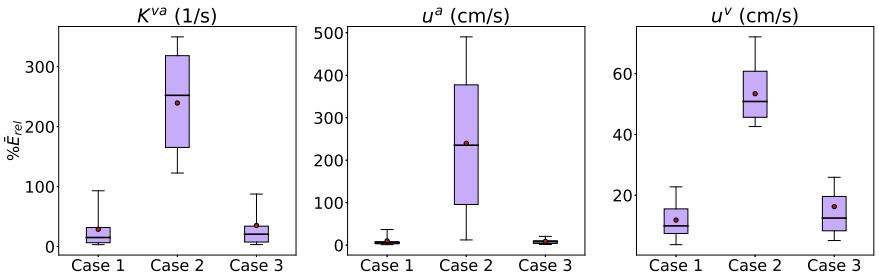

Figure 6.16: Boxplots illustrating the distribution of $\bar{E}_{rel}$ values for each 1D two-compartment case when applying noiseless measurement data. The whiskers denote the 5^{th}-95^{th} percentile, with the box region indicating the interquartile range. The medians are shown as the solid black line with the means shown as dark red points.

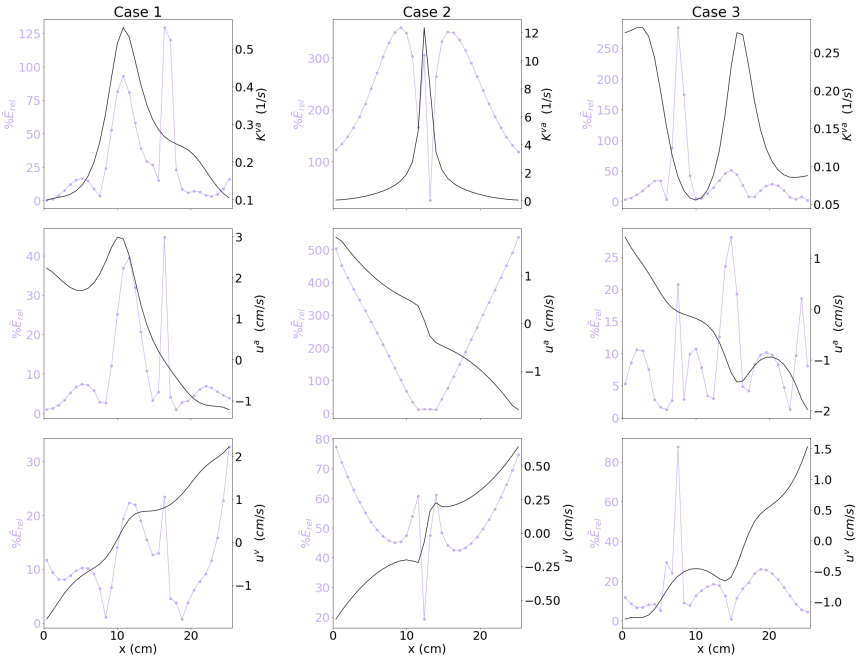

Figure 6.17: Spatial variation in $\bar{E}_{rel}$ values (purple) for each 1D two-compartment case compared to the ground truth values (black) when applying noiseless measurement data. To allow for comparison of different scales, the left-hand axis shows the $\bar{E}_{rel}$ range, while the right-hand axis shows the ground truth parameter range.

3 recovery of u^a, u^v and K^{va} achieved low $\bar{E}_{rel}$ values. In Case 1, u^a, u^v and K^{va} are recovered with average $\bar{E}_{rel}$ values of $10.1 \pm 11.9\%$, $11.85 \pm 7.1\%$, and $28.8 \pm 34.8\%$, respectively. Similarly, in Case 3, u^a, u^v and K^{va} are recovered with average $\bar{E}_{rel}$ values of $9.1 \pm 6.5\%$, $16.28 \pm 14.7\%$, and $35.1 \pm 54.9\%$, respectively. For Case 2 the recovery is unsuccessful, with large deviations shown between the recovery and ground truth parameter values (Figure 6.15). This is reflected in high $\bar{E}_{rel}$ values across all parameter groups of $239.9 \pm 162.9\%$, $53.4 \pm 11.6\%$, and $239.4 \pm 88.0\%$ for u^a, u^v and K^{va}, respectively. The mismatch in results between Case 2 and the other test cases creates a heavily biased distribution when looking at parameter recovery trends across all of the test cases. Therefore, each case is displayed separately when considering the parameter trends (Figure 6.16).

The tissue concentration reconstruction for the noiseless two-compartment cases (Figure 6.18) showed very low error when compared to the ground truth data. As illustrated in Figure 6.18 the network concentration and the ground truth are visually identical with the maximum error as a percentage of peak concentration below $\pm 1\%$, with the mean percentage difference below $\pm 0.2\%$.

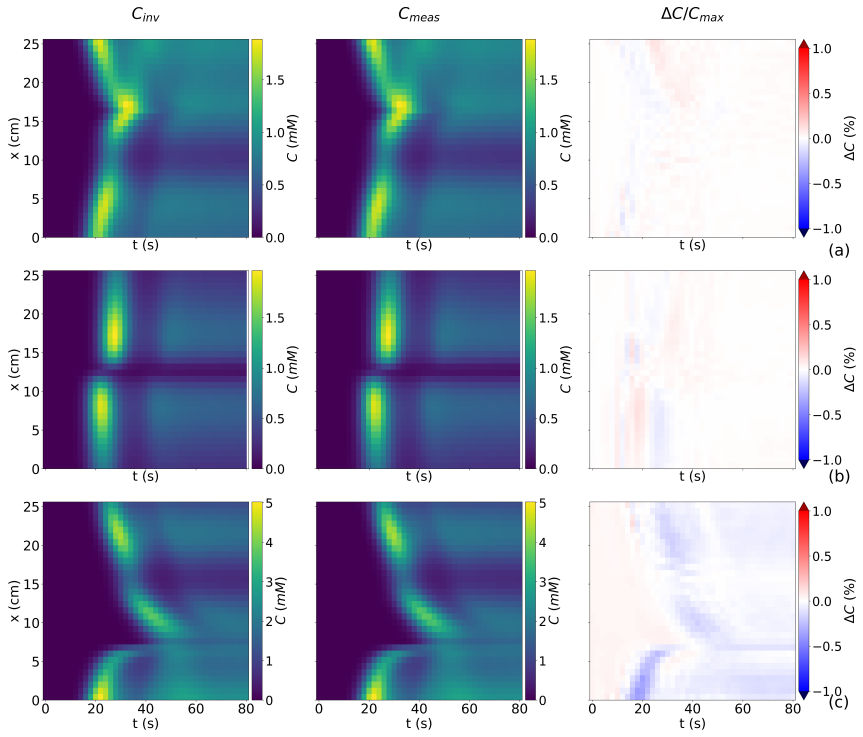

Figure 6.18: Tissue concentration evolution in tkPINNs recovery and ground truth systems for each 1D two-compartment case. First column shows C_{inv} recovered by network, second column shows ground truth C_{meas} with final column showing the difference between the two as a percentage of the maximal ground truth concentration. Cases 1, 2 and 3 are shown by rows (a) (b) and (c), respectively.

Measurement Noise

A summary of parameters retrieved from the tkPINNs method from systems with varying measurement SNR are shown in Figure 6.19. Variable agreement is shown between the ground truth parameter sets (black lines) and the retrieval at a variety of SNRs (coloured points). Broadly, across Cases 1 and 3 increased accuracy is seen for decreasing measurement noise. A detailed distribution of $\bar{E}_{rel}$ by each case and SNR is shown in Figure 6.20 for each parameter type.

At the highest SNR there is a marked decrease in parameter accuracy in comparison to the noiseless data recovery. At low SNR the parameter recovery appears to trend towards zero solutions, indicated in Figure 6.19 by the flattening of the parameter line close to zero, and also in Figure 6.20 by mean $\bar{E}_{rel}$ values close to 100%.

It is seen from Figures 6.19 and 6.20 Case 2 shows a poor parameter recovery with high $\bar{E}_{rel}$ values reported across all SNRs and parameter groups. A marked increase in $\bar{E}_{rel}$ is seen at SNR 20 for u^a and K^{va} parameters (Figure 6.20). Conversely, u^v parameter recovery shows stable $\bar{E}_{rel}$ values regardless of the SNR investigated.

Reduced Sampling

Parameter retrievals across each undersampling rate are detailed in Figure 6.21. A general trend of increasing parameter accuracy for higher time resolution measurements is shown. These trends are illustrated from the boxplots capturing the distribution of $\bar{E}_{rel}$ values reported in Figure 6.22.

Assessing the $\bar{E}_{rel}$ distributions (Figure 6.22), Case 2 showed high variability between different undersampling levels: a rise in $\bar{E}_{rel}$ between 2 and 6s followed by a subsequent fall for 8-10s. Conversely, Cases 1 and 3 showed $\bar{E}_{rel}$ values increasing under a decrease in sampling rate. The best and most consistently recovered parameter is u^a which has an $\bar{E}_{rel}$ range from 8.4-30.0% between sampling rates of 2-10s when across Cases 1 and 3.

Case 1 returns typically lower $\bar{E}_{rel}$ values in the presence of this undersampling across recovery for K^{va} and u^v than is shown in Case 3 (Figure 6.22). Across u^a,

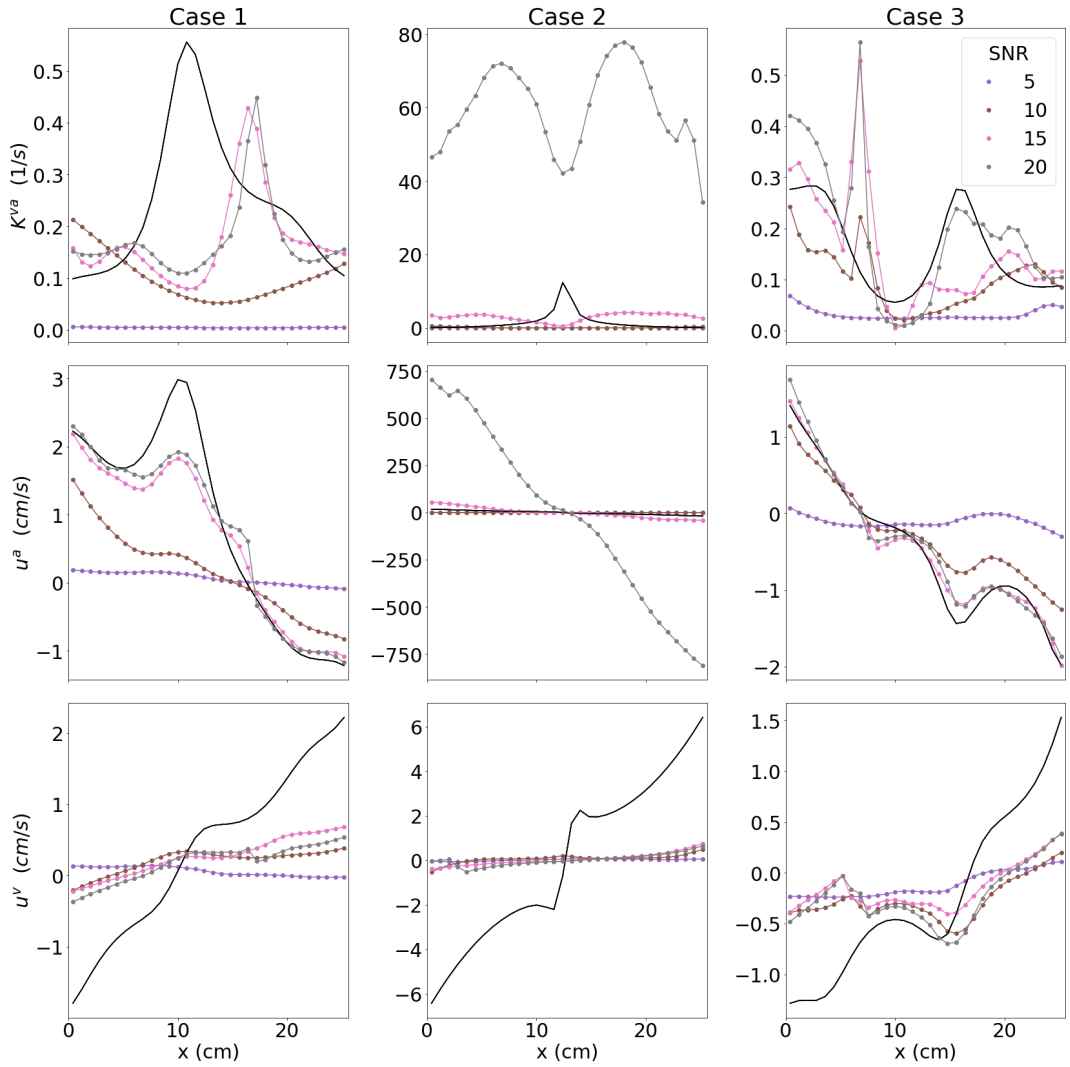

Figure 6.19: Parameter retrieval results from tkPINNs for 250,000 training iterations for measurement data with SNR 5, 10, 15 and 20 from each 1D two-compartment case. The noiseless case parameters are also displayed. Each noise level is displayed in a different colour.

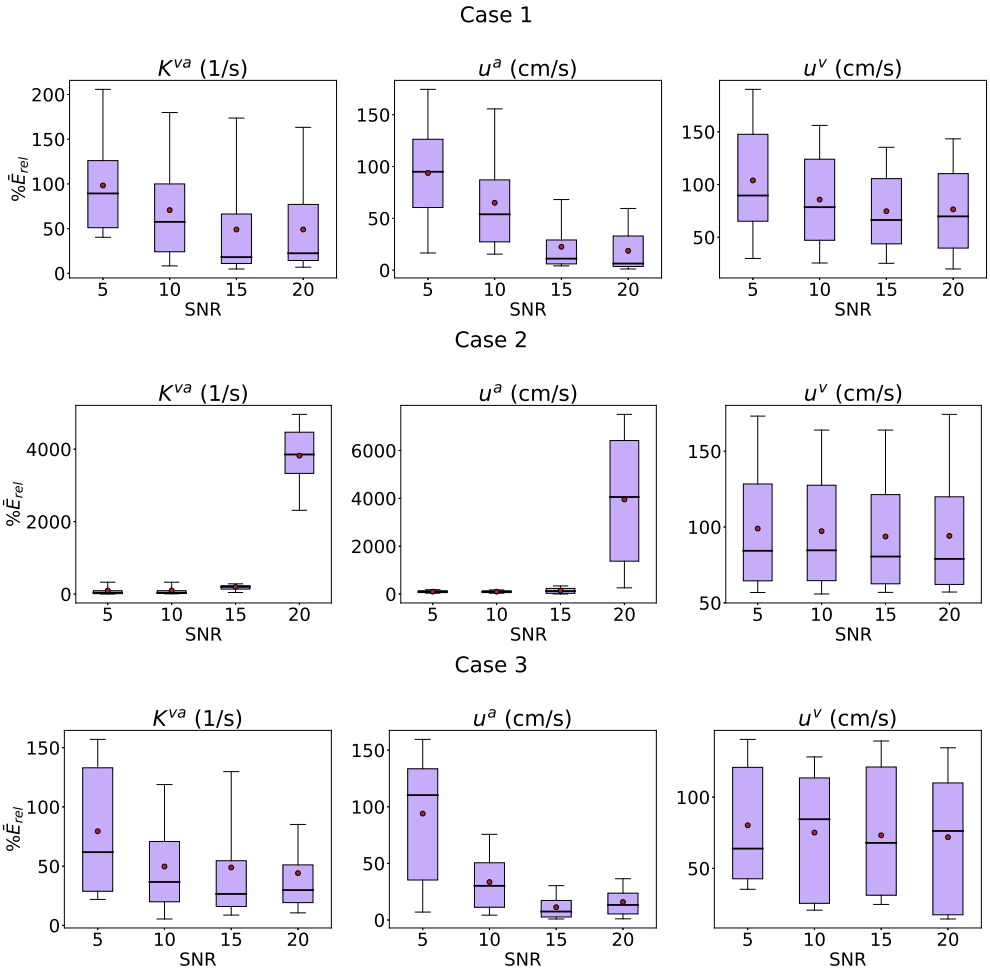

Figure 6.20: Distribution of each parameter error for each 1D two-compartment case relative to the absolute mean parameter value within each system. The distribution is shown across each measurement SNR. The whiskers denote the 5^{th}-95^{th} percentile, with the box region indicating the interquartile range. The medians are shown as the solid black line with the means shown as dark red points.

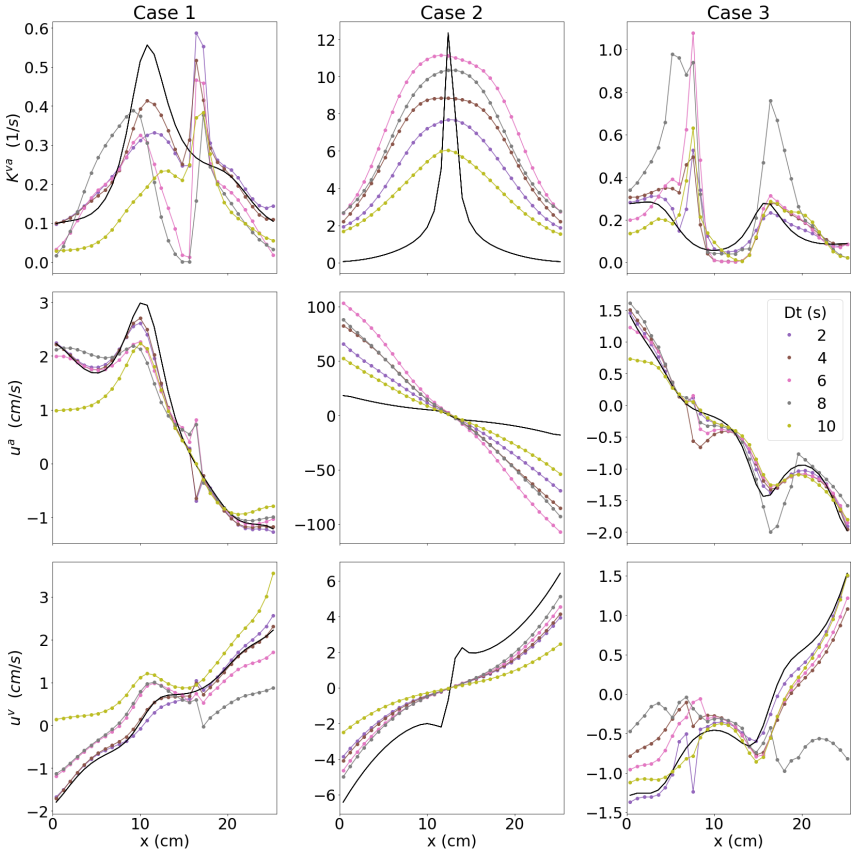

Figure 6.21: Parameter retrieval results from tkPINNs for 250,000 training iterations for measurement data with sampling resolution 2, 4, 6, 8 and 10s for each 1D two-compartment case. Each sampling level is displayed in a different colour.

the observed $\bar{E}_{rel}$ values for Cases 1 and 3 showed a similar range from 8.4-30.0% in Case 1 and 9.1-21.7% in Case 3.

6.3.4 One-Compartment in Two Dimensions

The resulting parameter retrieval is shown in Figure 6.24 using a 2D colourmap to denote velocity magnitude with a quiver plot overlaid to show the direction. The ground truth is shown with the colour bar and quiver arrows set to the same scale. The runtime for this two-compartment system in 2D was 88 hours.

Overall the recovery shows a promising reconstruction of the velocity pattern (Figure 6.24). This is also seen in the component-specific recoveries shown in Figure 6.25. Notably, the extreme areas in the velocity field are overestimated by the network, with high-velocity areas being overestimated while low-velocity areas are underestimated. Figure 6.26 indicates an inverse relationship between

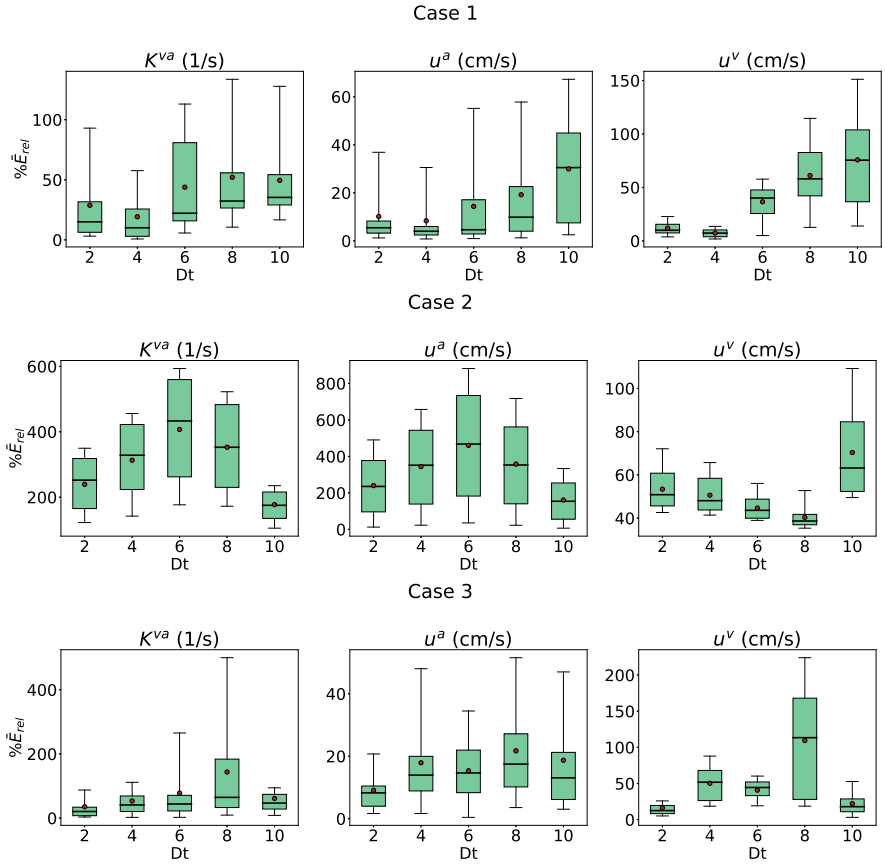

Figure 6.22: Distribution of each parameter error for each 1D two-compartment case relative to the absolute mean parameter value within each system. The distribution is shown across each undersampling rate. The whiskers denote the 5^{th}-95^{th} percentile, with the box region indicating the interquartile range. The medians are shown as the solid black line with the means shown as dark red points.

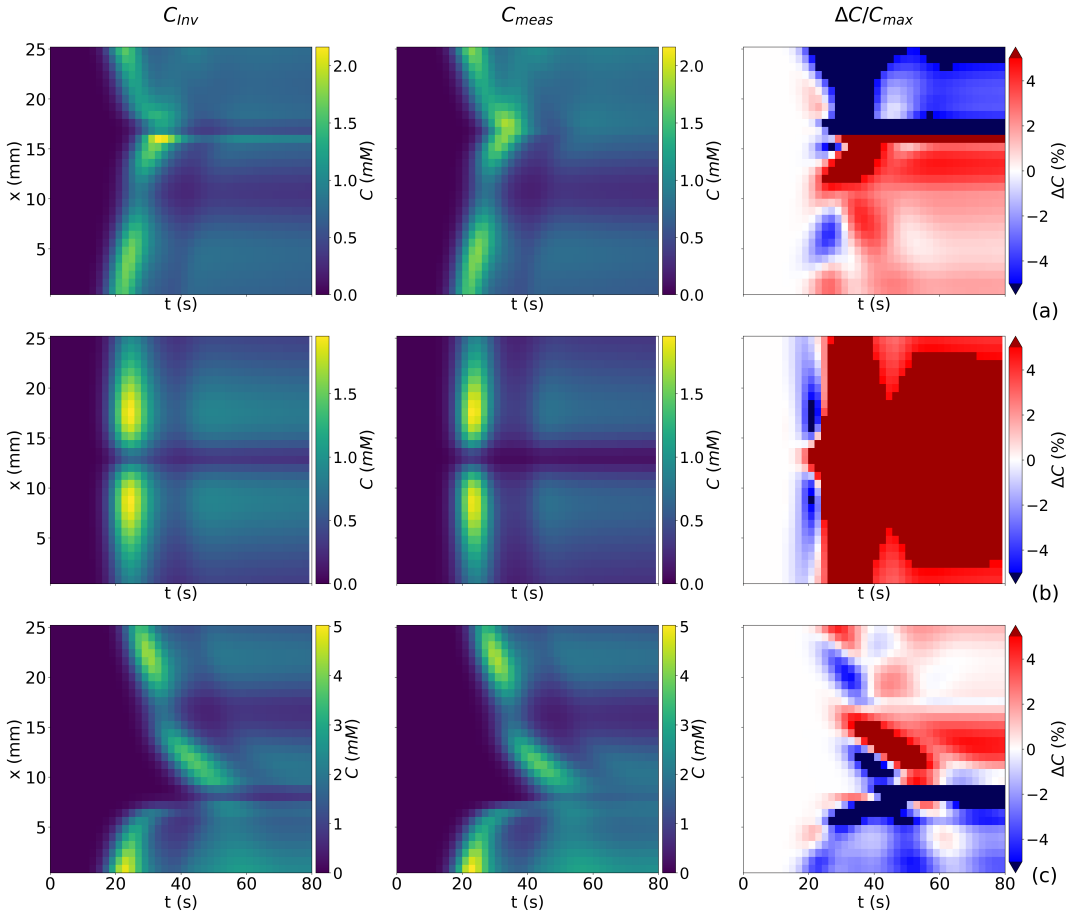

Figure 6.23: Tissue concentration evolution of tkPINNs network parameters using forwards model with true boundary influxes and ground truth systems for each 1D two-compartment case. First column shows C_{inv} from network parameters, second column shows ground truth C_{meas} with final column showing the difference between the two as a percentage of the maximal ground truth concentration. Cases 1, 2 and 3 are shown by rows (a) (b) and (c), respectively.

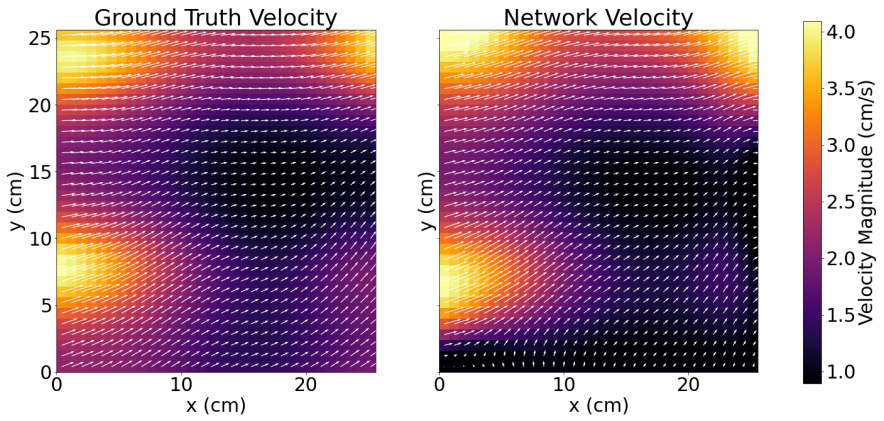

Figure 6.24: Overlaid colourmap and quiver plot showing the recovered velocity field for the one-compartment system in 2D with noiseless data. Left: network reconstruction and right: ground truth velocity.

the overall concentration seen by a voxel over the measurement time course and the error in the recovered velocity values at that point.

6.3.5 Loss Function Weighting Investigation

The resulting impact of small methodology adjustments are reported here. Alterations of weighting between the physics and data terms within the loss function is investigated. A full systematic evaluation of all tkPINNs network hyperparameters is out of the scope of this thesis. However, the following results shed light on an avenue that may be promising for future development.

Two alternative loss function weighting schemes are investigated. The resulting recovery for Case 1 at SNR 5 was assessed in both cases. The resulting parameter recoveries for data weighting of 1, 0.01, and 0.001 are shown in Figure 6.28. From these resulting parameter retrievals (Figure 6.28) it is seen that: u^a values are improved, K^{va} values are significantly reduced but remain inaccurate, and u^v values are largely unchanged. More aggressive down-weighting reversed this positive impact. Unfortunately, even the increased parameter accuracy from down-weighting the data by a factor of 0.01 does not produce comparable parameter recoveries for the noiseless cases.

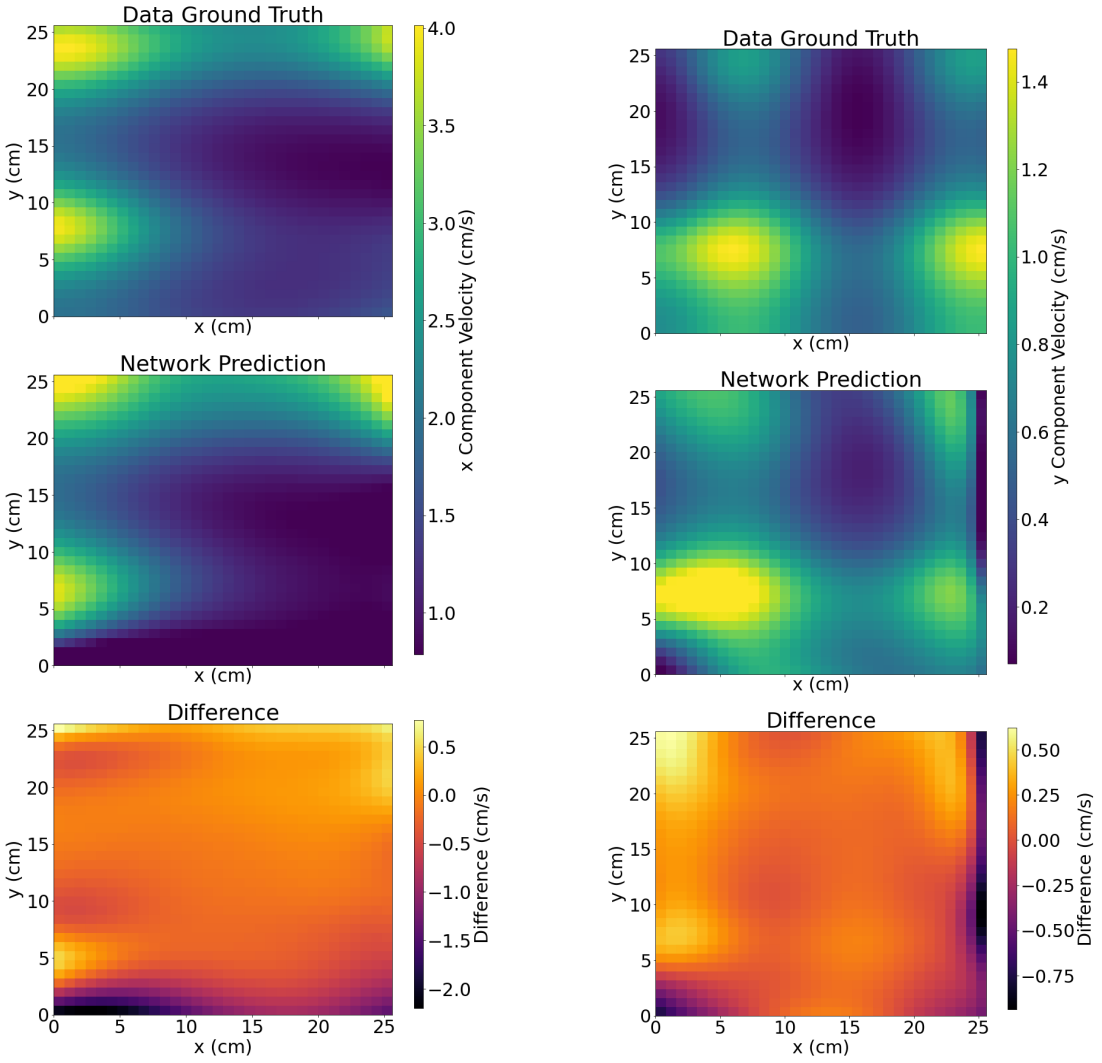

Figure 6.25: Velocity components for the two-compartment system in 2D. Left: x component velocity and, right: y component velocity. For each component, top: ground truth values, middle: network prediction and, bottom: error between network and ground truth values.

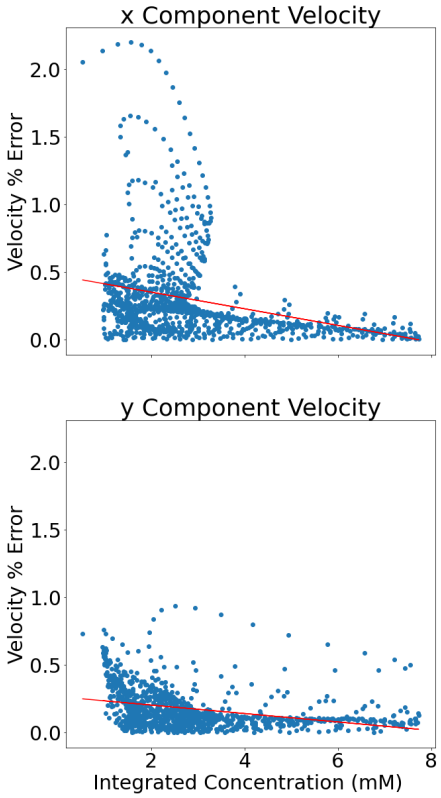

Figure 6.26: Correlation of x (top) and y (bottom) velocity component percentage errors for the two-compartment system in 2D compared against the integrated concentration in each voxel.

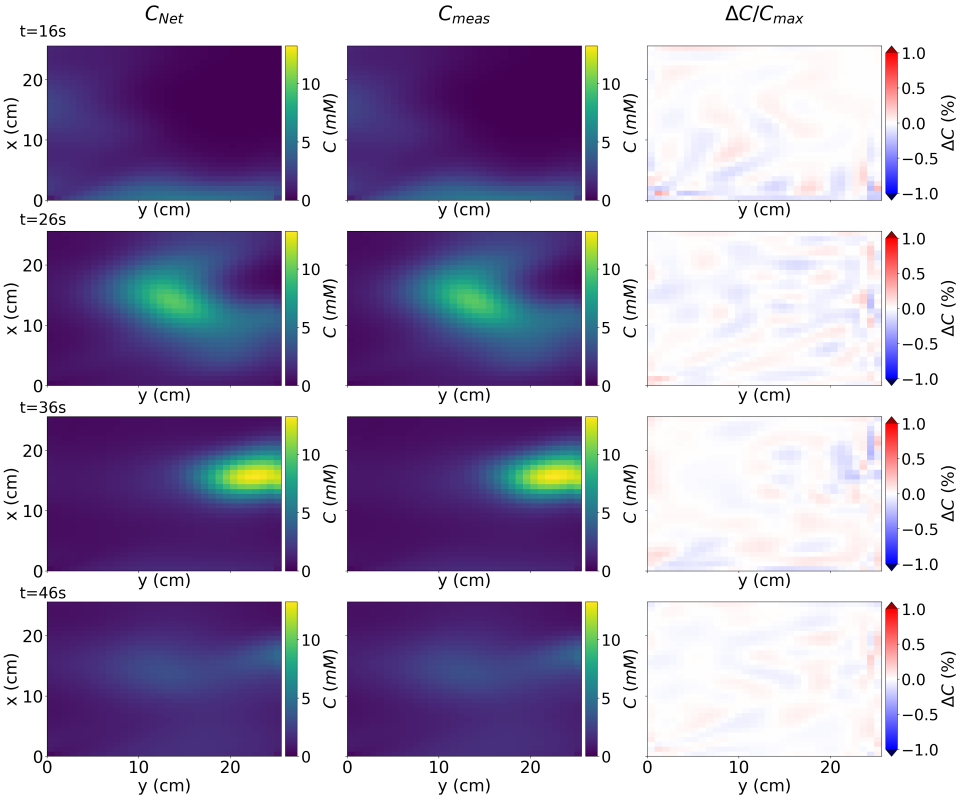

Figure 6.27: Tissue concentration evolution in tkPINNs recovery and ground truth systems for each 2D one-compartment case. First column shows C_{Net} recovered by network, second column shows ground truth C_{meas} with final column showing the difference between the two as a percentage of the maximal ground truth concentration. Each row shows the results from a different time point $t = 16, 26, 36, 46$s, top to bottom, respectively.

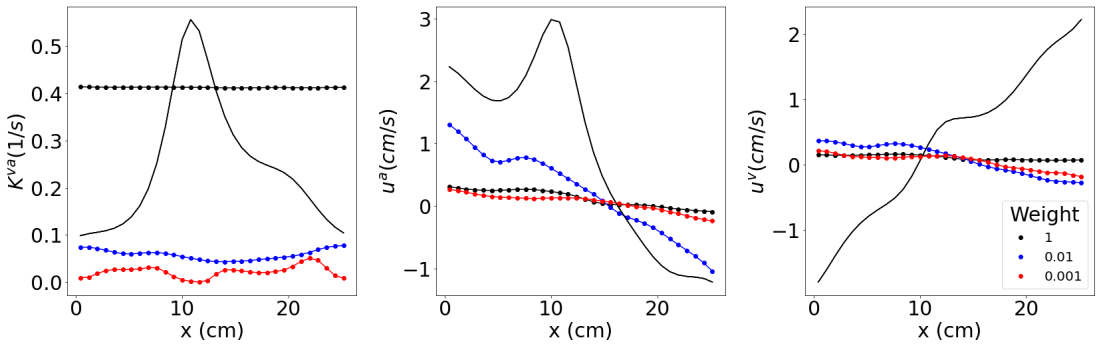

Figure 6.28: Parameter recovery for Case 1 of the two-compartment systems for different data weighting.

6.4 Discussion

The tkPINNs approach presented here allows for successful reconstruction of spatiotemporal parameters in some systems. One-compartment systems in both one and two dimensions showed low error parameter retrieval. Two-compartment systems gave more variable outcomes, particularly observed in Case 2. Across all system types, the concentration reconstructed from the network had a very low error. This close agreement to the measurement data, however, does not ensure a successful parameter-fitting outcome in two-compartment systems.

6.4.1 Hyperparameters

As outlined in Section 6.2.4, hyperparameters have to be chosen to implement a network. In the presented method, maximum iterations are chosen based on trends seen where parameter error appears to achieve a minimum. The loss function however (Figures 6.6 and 6.7) continued to decrease even though the parameter solutions were optimising away from the known solution. Ideally, exit conditions would be more data led. For example, estimation of SNR from measurement data could be applied to relax the data loss constraint. Additionally, relative parameter updates between iterations could also be used. Potential for this can be seen in Figure 6.3, where parameters remained consistent between iterations for optimal values. Implementation of such exit conditions, could be challenging due to randomised coordinate points used during training.

Other hyperparameters (summarised in Table 6.2) are chosen based on previous PINN implementations for a similar problem setup (Raissi *et al.*, 2020). There is no guarantee that these hyperparameters are the most optimised for spatiotemporal TK systems. While an indepth investigation into interplay and impact of hyperparameters was out of scope for this work, it is highlighted as a critical step needed in the future. A systematic study to assess the relative significance of hyperparameter variation on parameter accuracy is recommended for future work.

6.4.2 One-compartment

All one-compartment cases in 1D systems presented throughout this chapter showed low error parameter recoveries. This has been illustrated in the presence of increasing measurement noise and undersampling. The solutions produced indicate that these single-compartment systems are highly identifiable, which was previously observed in Chapter 5 with standard gradient descent optimisation. Additionally, several studies (Zhou *et al.*, 2021; Zhang *et al.*, 2022, 2023; Liu *et al.*, 2021a) covering similar one-compartment systems showed successful recovery, albeit with constraints placed on the spatial gradients of the velocity fields. An encouraging advancement presented within this chapter is recovery of such systems without the application of regularisation methods.

A preliminary extension of the network to higher spatial dimensions also provided encouraging results. The outputs achieved from this simple PINNs network setup for a one-compartment system in 2D showed that trends in velocity magnitude and direction are well recovered for the tested *in-silico* system (Figure 6.24). The increase in computing power needed to scale up to 2D using a full-sized batch is large even using this method, with one-compartment 1D systems taking ≈30 minutes, and one-compartment 2D systems taking ≈88 hours. Determination of an optimal batch size which decreases runtime, while preserving parameter accuracy, is a crucial step for future work.

6.4.3 Two-compartments

Parameter reconstructions are significantly worse within the two-compartment systems. Multiple solutions are seen across investigations where the network produces concentrations compatible with measurement data, and therefore a single optimal solution cannot be identified using only a goodness-of-fit assessment. The presence of multiple solutions is particularly seen when assessing data with measurement error (Figures 6.19 and 6.21), where recoveries produce entirely separate sets of solutions when compared to the known ground truth. The low sensitivity of the model to parameter perturbations and

covariance effects between model parameters as demonstrated in Chapter 5 could also contribute to these degenerate solutions.

Overall, the presence of noise appears a significant hindrance in the two-compartment systems for accurate parameter recovery solutions. The $\bar{E}_{rel}$ values associated with decreasing SNR are generally higher than $\bar{E}_{rel}$ values associated with decreased sampling resolution. This indicates that the method is not robust to noise and enforces strict requirements on appropriate data quality. As the method is shown to be more stable when temporal sampling is reduced, it may be appropriate to collect higher SNR data at a lower time resolution. This is a speculative recommendation as time undersampling still reduces accuracy of the retrieved parameters.

All two-compartment cases showed good agreement between the network and measurement concentrations when using noiseless fully sampled data (Figure 6.18). However, while the Case 2 network concentration shows very good agreement with measurement concentration (differences $< 1\%C_{max}$), the parameter recovery has a high $\bar{E}_{rel}$ (>200). To assess the accuracy of parameter solutions the original forwards model (Chapter 5) was applied showing a worse concentration reconstruction than the network output (Figure 6.23). As Cases 1 and 3 show close reconstruction to ground truth parameters most differences observed will be due to interpolation of transport parameters to the voxel interfaces. However, Case 2 showed large differences due to poorly recovered velocity values. This reinforces that the optimal solution cannot be identified from a close fit between measurement and network concentrations.

Boundary conditions are usually applied as additional data or physics loss terms in PINNs implementations. Unfortunately, boundary constraints are needed for separate compartments which are not directly accessible from measurement data. Lack of enforced boundary conditions is identified as a key issue within the tkPINNs setup. Adjustment of the network to allow estimation of compartmental influx could improve accuracy for other parameters. This was pursued during development, but a functional network was not realised due to unresolved compatibility issues with automatic differentiation. Alternatively, specifically designed systems with arterial-only regions at the boundaries could

be investigated, or regions with clear arterial input from scan data might be used.

Generally, poor reconstructions in the two-compartment systems have velocity in one compartment reconstructed more accurately than the other. While use of the velocity picture avoids soft constraint of flow incompressibility, definition in the flow picture could be useful. Recasting the network to use flows could leverage the interlinked changes in compartment blood flow, rather than the use of two independent velocity fields. However, with more parameter fields to solve, zero-value solutions (a common problem with PINNs) may become more prevalent.

An ever-present issue within PINNs are zero-value solutions. Networks can return zero-value solutions as by default they return 'perfect' physics loss terms. This hinges on the way the physics-based residual terms are calculated using automatic differentiation to determine if the physics fits at any point, rather than the standard method where physics is enforced via a full time course simulation. Constraints on the gradient of the physics loss term could be applied, which has been shown to improve identification of non-trivial solutions recovery in some PINNs applications (Leiteritz & Pflüger, 2021).

Alternative weighting strategies have been shown to improve outputs' convergence and accuracy from PINNs networks (Jagtap *et al.*, 2020). This is reflected in the presented results (Section 6.3.5) where the effect of down-weighting the data portion of the loss function by a factor of 0.01 and 0.001 is investigated.

6.4.4 General Observations

It is noteworthy to highlight differences between the internal working of the physics residuals calculated within the tkPINNs approach and the forwards model used to calculate the voxel-wise concentration evolution. As with any finite difference method forward model, the produced solution suffers from the effect of numerical diffusion (Kajishima *et al.*, 2017). Additionally, TKs describes the underlying mechanisms in an averaged sense rather than the exact underlying blood flow within blood vessels according to the full fluid dynamics

equations. Conversely, PINNs applies automatic differentiation at the collocation points with the system treated as continuous.

Recent forward modelling with PINNs (Krishnapriyan *et al.*, 2021) has identified that optimisation landscapes produced using soft constraints as a common failure mode which leads to poor solution identification. Alternative strategies, such as sequential training with temporal data showed error reduction of 1-2 orders of magnitude (Krishnapriyan *et al.*, 2021).

6.5 Conclusion

This chapter outlines a positive application of a PINNs-based method for the spatiotemporal retrieval of both 1D and 2D one-compartment systems. Across all tested 1D one-compartment systems the recovery appears robust to undersampling and increasing measurement noise. The presented tkPINNs method showed low error recovery with noiseless data for some two-compartment systems. This illustrates an encouraging preliminary application of PINNs methodology to the spatiotemporal TK problem. Multiple solutions are observed for two-compartment systems, indicating identifiability issues. Additional boundary condition constraints could help suppress multiple or incorrect solutions in two-compartment systems. Generally, hyperparameter optimisation is recommended to optimise the approach and reduce trivial solutions.

Chapter 7

Conclusions and Recommendations for Future Work

7.1 Conclusions

Here, the overall thesis conclusions are summarised and assessed in relation to the four main aims. Individual chapters contain detailed specific conclusions associated with the work presented within them.

Aim I: Evaluate K^{trans} quantification in current DCE-MRI perfusion analysis software.

Deviation in K^{trans} quantification is reported across 10 popular single voxel analysis pipelines currently available for DCE-MRI analysis (Chapter 3). This research highlights that there is some way to go in order to standardise K^{trans} quantification fully. However, due to the wide-scale uptake and considerable literature relating these PK quantities to a range of pathologies, these current techniques remain incredibly useful. Furthermore, the scalability of such methods cannot be overlooked due to the sheer amount of data present in DCE-MRI data sets.

Aim II: Review previous contributions of spatiotemporal modelling in perfusion MRI.

A set of 9 nested models of increasing complexity was identified within the

spatiotemporal literature review (Chapter 4). Implementations within the field were found to commonly target one-compartment systems with very sparse contributions for two-compartment systems. This work identified the need for new implementations, applying physically relevant assumptions, which target two-compartment systems. Such developments are important to appropriately model more complex biological systems.

Aim III: Investigate identifiability of one- and two-compartment spatiotemporal models.

For one-compartment systems this thesis illustrates good reconstruction of spatiotemporal parameters when considering data with and without measurement error (Chapters 5 and 6). This shows that these one-compartment systems can be readily characterised by their full spatial nature. When approaching fitting using standard optimisation strategies this research also details good reconstruction of concentration influx at the boundaries (Chapter 5).

For two-compartment perfusion systems this work highlights some identifiability issues in the form of multiple solutions (Chapters 5 and 6). When applying standard optimisation strategies in high SNR or noiseless data, concentration influx and velocity within the arterial compartment can be recovered well (Chapter 5).

Aim IV: Assess Physics Informed Neural Networks for suitability in scaling up spatiotemporal methods.

During work on standard optimisation for spatiotemporal systems, long computational runtimes are identified as a barrier for application to realistic datasets (Chapter 5). The potential suitability of a PINNs-based approach is demonstrated by several positive parameter reconstructions. A promising outcome for future application to realistic data volumes is the straightforward scalability of the method into 2D. It is proposed in this work that inability to properly integrate boundary conditions into the network causes additional identifiablity issues within the PINNs setup. Lack of boundary condition control reduces the suitability of the presented approach for multi-compartment

systems.

7.2 Overall Conclusions

Work within this thesis makes good progress towards addressing the multi-compartment problem by identifying concerns surrounding multiple solutions. Previous work in the field has focused on one-compartment systems or two-compartment systems with unrealistic assumptions to reduce the number of free parameters. The work on two-compartmental models with perfusion exchange (Chapters 5 and 6) provides an important addition to the field. This work has generated new open questions on fundamental system identifiability.

Spatiotemporal models show real promise as a paradigm change within perfusion MRI. However, further progression is needed to develop techniques appropriate for two-compartment models of full 4D DCE-MRI data sets. Additionally, parameters resulting from spatiotemporal approaches will require validation and benchmarking efforts before translation to any routine clinical setting is possible.

7.3 Recommendations for Future Work

To aid efforts in the standardisation of single voxel approaches, Chapter 3 offers several suggestions for future investigation. Questions that emerge about the origin of differences observed between software packages are critical to address in work going forward. For example, analysis of co-variance between model parameters and global input function influence are raised as an issue in Chapter 3. Further comparison of different software packages on single common datasets would be a useful approach to investigate this. Alternatively, individual software packages with variable pipeline setups could analyse differences produced by applying a variety of options to the same data set.

Work is required on increasing identifiability of two-compartment systems. This could be pursued via the modification of experimental conditions, to provide higher temporal resolution of data, or enhanced spatial information from multiple

bolus injections. Additionally, further development of the optimisation method is required, from underlying assumptions to implementation strategies. This could include fixing less important parameters to literature values, or appropriate use of regularisation to improve system identification.

Future work for the PINNs implementation should address scaling up the application to handle full 4D clinical DCE-MRI data, with 4D synthetic data produced for accuracy verification. Fundamentally, work on appropriate implementation of boundary conditions for PINNs based approaches is also required. This is especially crucial in multi-compartment systems where compartment concentrations are not accessible from measurement. Integration of system inputs as network produced quantities should be investigated. In terms of method design, the PINNs approach requires additional work on hyperparameter tuning, especially when fitting systems with appreciable measurement noise.

7.4 Summary

Overall this research has provided useful insights into future directions for perfusion imaging. It develops a spatiotemporal approach for one- and two-compartment systems from a standard gradient descent method to include machine learning techniques in the form of PINNs. Critically, it has detailed that two-compartment perfusion systems can have multiple possible solutions leading to identifiability issues. The PINNs approach is potentially suitable for transition into clinically relevant data sizes, although boundary conditions should be investigated in multi-compartment systems. This work sets the stage for future developments and identifies that further constraints on data quality and optimisation choices should be explored for more robust identification of two-compartment perfusion systems.

Appendix A

Additional Information for Chapter 3

Details contained in the supporting information from Shalom *et al.* (2024a) which were unsuitable for the main body of this thesis are include within this appendix.

A.1 RIDER NEURO Sequence Details

Content in this appendix section is included from Supporting Information 1 from Shalom *et al.* (2024a).

RIDER NEURO General Information:

- All MRI data sets were collected on the same 1.5T imaging magnet.
- All the patients within the full RIDER NEURO data set including the sub-set selected for the challenge have recurrent glioblastoma.
- For the test and retest visits, the images were taken 2 days apart, except for RIDER Neuro MRI-1086100996 which was not a part of the sub-set selected for the challenge or used for generation of either synthetic dataset.

The details of the main RIDER NEURO sequences are as follow:

- For T_1 mapping variable flip angle (VFA) 3D FLASH images in the axial plane were obtained using flip angles of 5°, 10°, 15°, 20°, 25°, and 30°TR of 4.43 ms, TE of 2.1 ms, 2 signal averages.

- DCE-MR images acquired using a 3D FLASH technique in the axial plane were obtained during the intravenous injection of 0.1 mmol/kg of magnevist at 3 mL/second, started at 24 seconds after the scan had begun. The dynamic images were obtained with a flip angle (FA) of 25°, TR of 3.8 ms, TE of 1.8 ms using a 1x1x5 mm3 voxel size. The 16-slice volume was obtained every 4.8 sec (temporal resolution).

Accompanying data:

- Contrast-enhanced 3D FLASH: The patients underwent whole brain 3D FLASH imaging in the sagittal plane after the administration of magnevist with the following parameters: TR= 8.6 ms, TE= 4.1 ms, FA= 20°, number of averages= 1, matrix size= 256x256, voxel size= 1 mm3.

- Contrast-enhanced 3D FLAIR: The patients have 3D FLAIR sequences in the sagittal plane after the administration of magnevist with the following parameters: TR= 6000 ms, TE= 353 ms, TI= 2200 ms, FA= 180°, number of average= 1, matrix size= 256x256; voxel size = 1 mm3.

For additional information, see the RIDER NEURO homepage (https://wiki.cancerimagingarchive.net/display/Public/RIDER+NEURO+MRI).

A.2 Supplementary Data Tables

Data tables included in this appendix section are from Supporting Information 2, 3 and 4 from Shalom *et al.* (2024a).

Table A.1: Recovered K^{trans} values (min^{-1}) over the tumour mask for each clinical visit in all entries. Here C_1v_1 denotes clinical patient set 1 at visit 1, the naming system follows directly for the remaining sets.

Visit	ROCKET SHIP	ALICE	ImageJ/ MRIcron	MRI-QAMPER	OHSU	Madym	DCE-NET	FireVoxel	PerfLab	UW QBI lab
C_1v_1	6.4E-2	1.1E-1	2.9E-2	2.0E-1	5.2E-2	3.4E-2	2.7E-2	9.1E-2	2.5E-2	7.2E-2
C_1v_2	6.2E-2	1.3E-1	1.4E-2	1.4E-1	1.1E-1	6.2E-2	4.9E-2	1.8E-1	6.1E-2	1.4E-1
C_2v_1	1.2E-2	1.0E-1	1.0E-1	7.6E-2	7.9E-2	4.7E-2	3.7E-2	9.0E-2	3.8E-2	9.6E-2
C_2v_2	5.2E-2	3.6E-1	7.6E-2	1.8E-1	3.9E-2	3.0E-2	2.3E-2	6.4E-2	2.9E-2	5.2E-2
C_3v_1	1.2E-2	3.2E-1	3.1E-3	1.9E-1	1.5E-1	6.3E-2	5.0E-2	3.1E-1	3.5E-2	1.1E-1
C_3v_2	6.2E-3	2.8E-1	1.6E-2	2.0E-1	6.1E-2	6.1E-2	4.9E-2	1.6E-1	6.6E-2	1.0E-1
C_4v_1	7.0E-2	1.8E-1	2.2E-3	8.1E-2	4.8E-2	3.2E-2	2.3E-2	1.3E-1	3.0E-2	7.1E-2
C_4v_2	1.3E-2	1.6E-1	1.3E-2	1.6E-1	7.9E-2	3.0E-2	2.2E-2	1.7E-1	2.1E-2	7.3E-2
C_5v_1	4.6E-2	1.2E-1	4.2E-3	1.7E-1	7.3E-2	4.5E-2	3.2E-2	1.6E-1	2.7E-2	1.0E-1
C_5v_2	3.2E-2	1.9E-1	2.5E-2	1.3E-1	5.8E-2	4.3E-2	3.3E-2	1.8E-1	2.5E-2	8.9E-2
C_6v_1	1.2E-2	2.0E-1	4.4E-2	2.3E-1	5.4E-2	6.0E-2	3.2E-2	3.3E-1	5.6E-2	1.9E-1
C_6v_2	5.8E-3	1.8E-1	2.5E-2	2.2E-1	3.9E-2	2.5E-2	1.7E-2	1.1E-1	2.0E-2	8.0E-2
C_7v_1	7.6E-2	2.8E-1	4.0E-2	4.0E-1	5.8E-2	3.8E-2	2.5E-2	1.1E-1	4.5E-2	7.6E-2
C_7v_2	1.6E-2	2.8E-1	4.2E-2	3.4E-1	9.7E-2	6.2E-2	4.9E-2	2.3E-1	3.0E-2	2.0E-1
C_8v_1	2.3E-2	2.4E-1	3.0E-2	2.5E-1	5.5E-2	3.1E-2	2.1E-2	1.3E-1	1.7E-2	7.7E-2
C_8v_2	2.2E-1	3.5E-1	4.3E-2	2.2E-1	4.9E-2	3.4E-2	1.7E-2	1.0E-1	2.0E-2	7.0E-2

Table A.2: Recovered K^{trans} values (min^{-1}) over the tumour mask for each synthetic visit in all entries and the ground truth DRO. Here S_1v_1 denotes synthetic patient 1 at visit 1, the naming system follows directly the remaining sets.

Visit	DRO	ROCKET SHIP	ALICE	ImageJ/ MRIcron	MRI-QAMPER	OHSU	Madym	DCE-NET	Fire Voxel	Perf Lab	UW QBI lab
S_1v_1	1.1E-2	5.7E-7	3.8E-1	1.3E-2	2.1E-2	2.5E-2	1.7E-2	1.4E-2	3.7E-2	2.0E-2	3.6E-2
S_1v_2	4.2E-3	1.1E-5	8.8E-2	6.0E-3	3.7E-2	1.2E-2	7.5E-3	5.9E-3	2.2E-1	1.1E-2	1.8E-2
S_2v_1	1.1E-1	1.3E-1	1.8E-1	7.1E-2	1.0E-1	2.1E-1	1.2E-1	1.1E-1	3.5E-1	1.0E-1	2.6E-1
S_2v_1	6.9E-2	6.2E-2	1.6E-1	4.8E-2	1.0E-1	1.7E-1	8.2E-2	7.3E-2	2.0E-1	1.1E-1	1.7E-1

Table A.3: Values of the recovered K^{trans} values (min^{-1}) over the tumour mask for each visit from the neutral evaluators in all reproduced submissions. Here C_1v_1 and S_1v_1 denote clinical and synthetic patient 1 at visit 1, respectively. The naming system follows directly for the remaining sets.

Visit	ROCKET SHIP	MRI-QAMPER	Maydm	DCE-NET	FireVoxel	PerfLab
C_1v_1	1.0E-1	1.9E-1	3.4E-2	2.7E-2	8.6E-5	2.5E-2
C_1v_2	4.1E-2	1.4E-1	6.2E-2	4.9E-2	1.2E-2	6.0E-2
C_2v_1	1.4E-2	7.6E-2	4.7E-2	3.6E-2	7.0E-2	3.7E-2
C_2v_2	3.7E-2	1.8E-1	3.0E-2	2.3E-2	6.0E-2	2.7E-2
C_3v_1	2.6E-2	1.9E-1	6.3E-2	5.0E-2	9.5E-2	3.4E-2
C_3v_2	1.3E-2	2.0E-1	6.1E-2	4.9E-2	5.7E-2	5.8E-2
C_4v_1	2.5E-2	8.0E-2	3.2E-2	2.3E-2	1.2E-1	2.9E-2
C_4v_2	2.5E-2	1.6E-1	3.0E-2	2.2E-2	1.3E-1	2.0E-2
C_5v_1	2.6E-2	1.7E-1	4.5E-2	3.2E-2	9.7E-2	2.9E-2
C_5v_2	2.8E-2	1.4E-1	4.3E-2	3.3E-2	8.8E-2	2.6E-2
C_6v_1	2.2E-2	2.3E-1	6.0E-2	3.2E-2	3.5E-1	5.4E-2
C_6v_2	2.0E-2	2.5E-1	2.5E-2	1.7E-2	2.9E-1	1.9E-2
C_7v_1	4.9E-2	4.2E-1	3.8E-2	2.5E-2	1.7E-1	4.4E-2
C_7v_2	5.9E-2	3.2E-1	6.2E-2	4.9E-2	6.0E-3	3.2E-2
C_8v_1	3.1E-2	2.5E-1	3.1E-2	2.1E-2	6.2E-2	1.8E-2
C_8v_2	2.6E-2	2.2E-1	3.4E-2	1.7E-2	1.4E-1	2.2E-2
S_1v_1	3.2E-7	8.6E-2	1.7E-2	1.4E-2	1.6E-2	2.0E-2
S_1v_2	8.0E-6	2.2E-1	7.5E-3	5.5E-3	5.4E-3	1.1E-2
S_2v_1	1.6E-1	9.2E-2	1.2E-1	1.1E-1	4.2E-4	1.0E-1
S_2v_2	1.2E-1	1.0E-1	8.2E-2	7.4E-2	4.4E-2	1.1E-1

Table A.4: A summary of OSIPI scores for all entries. 95% confidence intervals for $Score_{accuracy}$, $Score_{repeat}$, and $Score_{reproduce}$ are shown with ± notation. Confidence intervals are generated using scores for each $\sigma\mu$ term in the summations (Table 3.3) separately.

Submission	Rank	$Score_{accuracy}$	$Score_{repeat}$	$Score_{reproduce}$	$OSIPI_{silver}$ (%)	$OSIPI_{gold}$ (%)
DCE-NET	1	0.92 ± 0.06	0.85 ± 0.08	1.00 ± 0.00	78	78
Madym	2	0.85 ± 0.07	0.85 ± 0.08	1.00 ± 0.00	73	73
PerfLab	3	0.78 ± 0.11	0.80 ± 0.08	0.98 ± 0.01	62	61
MRI-QAMPER	4	0.72 ± 0.18	0.86 ± 0.08	0.93 ± 0.06	62	57
FireVoxel	5	0.57 ± 0.04	0.78 ± 0.07	0.65 ± 0.09	45	29
ROCKET SHIP	6	0.59 ± 0.28	0.64 ± 0.12	0.74 ± 0.05	37	28
ImageJ/ MRIcron	7	0.85 ± 0.04	0.68 ± 0.12	N/R	58	N/R
OHSU	8	0.67 ± 0.04	0.79 ± 0.07	N/R	53	N/R
UW QIBlab	9	0.61 ± 0.05	0.81 ± 0.09	N/R	50	N/R
ALICE	10	0.54 ± 0.17	0.86 ± 0.09	N/R	46	N/R

Appendix B

Data Tables for Chapter 5

Data tables in the following appendix detail the mean (μ) and standard deviation (σ) of E_{rel} values from parameter recoveries in Chapter 5.

B.1 One-Compartment Systems

Table B.1: E_{rel} values for one-compartment system across all cases in noiseless data using the standard optimisation approach.

Case	E_{rel}			
	u		J	
	μ	σ	μ	σ
1	0.26	0.20	2.51	4.31
2	0.38	0.25	3.13	4.93
3	0.48	0.37	3.01	4.92
Overall	0.37	0.30	2.88	4.73

Table B.2: E_{rel} values for one-compartment system Case 1 for each initial guess using the standard optimisation approach.

Case	Guess	E_{rel}			
		u		J	
		μ	σ	μ	σ
1	1	0.26	0.20	2.53	4.29
1	2	0.27	0.20	2.53	4.29
1	3	0.26	0.20	2.48	4.29
1	4	0.26	0.20	2.49	4.33
1	5	0.26	0.20	2.50	4.33
1	6	0.26	0.20	2.52	4.32

!

Table B.3: E_{rel} values for one-compartment system Case 2 for each initial guess using the standard optimisation approach.

Case	Guess	E_{rel}			
		u		J	
		μ	σ	μ	σ
2	1	0.38	0.25	3.09	4.95
2	2	0.39	0.26	3.18	4.91
2	3	0.40	0.27	3.20	4.91
2	4	0.38	0.25	3.14	4.91
2	5	0.37	0.24	3.06	4.94
2	6	0.38	0.26	3.11	4.95

Table B.4: E_{rel} values for one-compartment system Case 3 for each initial guess using the standard optimisation approach.

Case	Guess	E_{rel}			
		u		J	
		μ	σ	μ	σ
3	1	0.47	0.37	3.01	4.92
3	2	0.48	0.37	2.99	4.94
3	3	0.49	0.38	3.06	4.90
3	4	0.47	0.36	2.99	4.92
3	5	0.48	0.37	3.02	4.91
3	6	0.48	0.37	3.01	4.91

B.2 Two-Compartment Systems

Table B.5: E_{rel} values for all one-compartment system cases across each SNR level using the standard optimisation approach.

SNR	E_{rel}			
	u		J	
	μ	σ	μ	σ
5	2.42	2.78	16.08	15.57
10	1.15	1.26	8.54	8.34
15	0.82	0.90	6.47	6.71
20	0.67	0.66	5.18	5.38

Table B.6: E_{rel} values for one-compartment system Case 1 across each SNR level using the standard optimisation approach.

Case	SNR	E_{rel}			
		u		J	
		μ	σ	μ	σ
1	5	2.26	2.84	11.51	10.78
1	10	1.22	1.32	6.43	6.03
1	15	0.81	0.82	5.52	5.01
1	20	0.66	0.69	4.40	4.51

Table B.7: E_{rel} values for one-compartment system Case 2 across each SNR level using the standard optimisation approach.

Case	SNR	E_{rel}			
		u		J	
		μ	σ	μ	σ
2	5	2.70	3.22	19.57	18.76
2	10	1.13	1.46	10.14	9.26
2	15	0.80	0.99	7.95	8.06
2	20	0.65	0.63	6.11	6.11

Table B.8: E_{rel} values for one-compartment system Case 3 across each SNR level using the standard optimisation approach.

Case	SNR	E_{rel}			
		u		J	
		μ	σ	μ	σ
3	5	2.29	2.17	17.15	14.99
3	10	1.10	0.95	9.05	8.90
3	15	0.85	0.90	5.94	6.44
3	20	0.70	0.67	5.02	5.27

Table B.9: E_{rel} values for all one-compartment system cases across each undersampling level using the standard optimisation approach.

Dt	E_{rel}			
	u		J	
	μ	σ	μ	σ
2	0.37	0.29	2.85	4.74
4	0.43	0.54	7.60	11.97
6	2.15	1.97	12.65	16.20
8	14.40	15.28	24.64	47.34

Table B.10: E_{rel} values for one-compartment system Case 1 across each undersampling level using the standard optimisation approach.

Case	Dt	E_{rel}			
		u		J	
		μ	σ	μ	σ
1	2	0.26	0.19	2.48	4.30
1	4	0.34	0.46	6.60	11.29
1	6	4.70	0.79	12.13	14.18
1	8	33.19	12.35	55.23	72.26

Table B.11: E_{rel} values for one-compartment system Case 2 across each undersampling level using the standard optimisation approach.

Case	Dt	E_{rel}			
		u		J	
		μ	σ	μ	σ
2	2	0.32	0.23	2.79	4.65
2	4	0.38	0.40	7.42	11.88
2	6	2.61	2.19	12.70	15.83
2	8	18.50	17.15	31.87	56.37

Table B.12: E_{rel} values for one-compartment system Case 3 across each undersampling level using the standard optimisation approach.

Case	Dt	E_{rel}			
		u		J	
		μ	σ	μ	σ
3	2	0.37	0.29	2.85	4.74
3	4	0.43	0.54	7.60	11.97
3	6	2.15	1.97	12.65	16.20
3	8	14.40	15.28	24.64	47.34

Table B.13: E_{rel} values for all two-compartment system cases across all initial guesses.

Parameter	Erel	
	mean	std
Kva	44.65	76.52
ua	27.06	82.04
uv	54.87	121.37
Ja	12.99	38.05

Table B.14: E_{rel} values for two-compartment system Case 1 for each initial guess.

Case	Guess	E_{rel}							
		u^a		u^v		K^{va}		J^a	
		μ	σ	μ	σ	μ	σ	μ	σ
1	1	6.28	7.98	3.19	4.50	23.84	27.96	4.26	4.19
1	2	6.08	7.78	3.30	4.67	22.72	28.70	4.38	4.22
1	3	8.34	9.01	16.69	15.10	35.86	37.67	4.44	3.97
1	4	19.11	25.82	273.54	186.44	77.69	72.13	12.71	12.25
1	5	19.04	25.80	273.73	186.68	77.17	72.55	12.62	12.20
	Total	11.77	18.54	114.09	175.98	47.46	57.60	7.68	9.31

Table B.15: E_{rel} values for two-compartment system Case 2 for each initial guess.

Case	Guess	E_{rel}							
		u^a		u^v		K^{va}		J^a	
		μ	σ	μ	σ	μ	σ	μ	σ
2	1	9.08	9.14	6.15	5.42	32.32	50.32	4.92	3.04
2	2	9.11	9.03	6.77	6.15	47.67	128.49	4.93	3.10
2	3	11.04	17.71	7.30	8.02	40.24	79.14	5.01	4.06
2	4	274.88	177.86	157.55	136.38	115.51	165.35	93.54	117.93
2	5	15.64	29.55	7.64	9.43	40.80	69.00	5.24	5.58
	Total	63.95	133.14	37.08	85.98	55.31	111.41	22.73	63.63

Table B.16: E_{rel} values for two-compartment system Case 3 for each initial guess.

Case	Guess	E_{rel}							
		u^a		u^v		K^{va}		J^a	
		μ	σ	μ	σ	μ	σ	μ	σ
3	1	5.42	7.54	13.94	17.99	34.66	38.42	8.36	8.15
3	2	5.41	7.71	13.09	16.53	30.31	39.50	8.60	8.16
3	3	5.46	7.66	13.02	16.32	30.37	39.27	8.60	8.10
3	4	5.51	7.62	13.21	16.60	30.33	39.27	8.55	8.03
3	5	5.44	7.79	13.99	18.32	30.32	39.18	8.70	8.33
	Total	5.45	7.66	13.45	17.18	31.20	39.17	8.56	8.16

Table B.17: E_{rel} values for all two-compartment systems at each SNR.

SNR	E_{rel}							
	u^a		u^v		K^{va}		J^a	
	μ	σ	μ	σ	μ	σ	μ	σ
5	49.79	95.26	83.30	158.33	95.16	151.94	37.03	23.26
10	35.72	72.77	32.16	47.15	89.47	184.74	2.53	1.86
15	27.94	63.00	23.57	47.20	69.13	114.24	11.55	6.41
20	22.47	50.92	24.06	52.26	60.44	74.40	6.59	2.38

Table B.18: E_{rel} values for two-compartment system Case 1 at each SNR.

Case	SNR	E_{rel}							
		u^a		u^v		K^{va}		J^a	
		μ	σ	μ	σ	μ	σ	μ	σ
1	5	19.76	20.06	123.05	176.07	83.35	69.81	33.70	32.54
1	10	15.62	17.64	27.23	28.94	74.56	63.68	17.69	16.19
1	15	11.82	12.84	22.10	24.23	52.47	49.22	10.76	10.34
1	20	10.27	13.73	11.58	14.08	48.37	46.72	8.96	7.40

Table B.19: E_{rel} values for two-compartment system Case 2 at each SNR.

Case	SNR	E_{rel}							
		u^a		u^v		K^{va}		J^a	
		μ	σ	μ	σ	μ	σ	μ	σ
2	5	115.68	141.69	95.85	196.87	130.83	242.84	63.34	71.74
2	10	80.14	111.53	44.79	70.73	137.56	302.24	40.00	51.88
2	15	63.15	98.71	26.96	72.46	99.63	181.11	14.65	14.77
2	20	48.91	80.14	32.63	80.37	83.79	104.74	20.46	22.42

Table B.20: E_{rel} values for two-compartment system Case 3 at each SNR.

Case	SNR	E_{rel}							
		u^a		u^v		K^{va}		J^a	
		μ	σ	μ	σ	μ	σ	μ	σ
3	5	13.95	14.79	30.98	31.13	71.29	58.59	23.94	20.64
3	10	11.38	12.93	24.46	24.18	56.30	57.84	14.98	12.29
3	15	8.84	11.65	21.65	28.80	55.30	50.32	10.59	9.60
3	20	8.24	10.87	27.97	35.92	49.16	51.36	10.33	9.16

Table B.21: E_{rel} values for all two-compartment systems at each undersampling rate using the standard optimisation approach.

Dt	E_{rel}							
	u^a		u^v		K^{va}		J^a	
	μ	σ	μ	σ	μ	σ	μ	σ
2	30.27	40.23	6.93	8.40	7.76	12.04	6.78	5.79
4	67.85	94.47	14.42	16.97	220.43	371.22	16.27	13.91
6	73.07	62.93	33.96	33.84	138.48	178.73	34.82	29.11
8	81.68	110.28	53.18	46.22	159.43	181.11	67.97	50.68

Table B.22: E_{rel} values for two-compartment system Case 1 at each undersampling rate.

Case	Dt	E_{rel}							
		u^a		u^v		K^{va}		J^a	
		μ	σ	μ	σ	μ	σ	μ	σ
1	2	6.28	7.98	3.19	4.50	23.84	27.96	4.26	4.19
1	4	13.10	8.75	595.99	446.69	69.63	41.84	20.26	16.75
1	6	30.80	33.47	329.84	188.00	92.09	80.26	23.89	22.33
1	8	41.42	41.49	344.64	132.39	56.26	44.51	56.64	50.49

Table B.23: E_{rel} values for two-compartment system Case 2 at each undersampling rate using the standard optimisation approach.

Case	Dt	E_{rel}							
		u^a		u^v		K^{va}		J^a	
		μ	σ	μ	σ	μ	σ	μ	σ
2	2	9.08	9.14	6.15	5.42	32.32	50.32	4.92	3.04
2	4	20.74	25.05	26.96	24.54	73.49	139.59	13.91	13.20
2	6	57.29	33.52	47.72	38.90	48.44	44.05	47.20	37.21
2	8	78.53	38.05	110.41	158.74	94.89	167.19	59.72	49.92

Table B.24: E_{rel} values for two-compartment system Case 3 at each undersampling rate using the standard optimisation approach.

Case	Dt	E_{rel}							
		u^a		u^v		K^{va}		J^a	
		μ	σ	μ	σ	μ	σ	μ	σ
3	2	5.42	7.54	13.94	17.99	34.66	38.42	8.36	8.15
3	4	9.43	9.68	38.33	40.45	60.41	73.82	9.49	8.98
3	6	13.80	15.22	37.87	63.24	78.69	49.98	20.30	17.67
3	8	39.59	47.67	23.25	20.73	93.88	74.72	42.24	56.33

Appendix C

Data Tables for Chapter 6

Data tables in the following appendix detail the mean (μ) and standard deviation (σ) of E_{rel} values from parameter recoveries in Chapter 6.

C.1 One-Compartment Systems

Table C.1: E_{rel} values assessed across one-compartment cases for noiseless data using tkPINNs method.

Case	E_{rel}	
	u	
	μ	σ
1	4.28	4.41
2	6.68	5.46
3	4.97	3.58

Table C.2: E_{rel} values assessed across all one-compartment cases at each SNR level.

SNR	E_{rel}	
	u	
	μ	σ
5	6.61	5.25
10	5.36	4.58
15	5.27	4.67
20	5.25	4.68

C.2 Two-Compartment Systems

Table C.3: E_{rel} values assessed across one-compartment Case 1 at each SNR level using tkPINNs method.

Case	SNR	E_{rel}	
		u	
		μ	σ
1	5	7.50	6.52
1	10	4.79	4.58
1	15	4.25	4.41
1	20	4.15	4.31

Table C.4: E_{rel} values assessed across one-compartment Case 2 at each SNR level using tkPINNs method.

Case	SNR	E_{rel}	
		u	
		μ	σ
2	5	6.23	4.49
2	10	6.29	5.19
2	15	6.48	5.40
2	20	6.52	5.49

Table C.5: E_{rel} values assessed across one-compartment Case 3 at each SNR level using tkPINNs method.

Case	SNR	E_{rel}	
		u	
		μ	σ
3	5	6.10	4.33
3	10	5.02	3.70
3	15	5.08	3.79
3	20	5.09	3.76

Table C.6: E_{rel} values assessed across all one-compartment at each undersampling level using tkPINNs method.

Dt	E_{rel}	
	u	
	μ	σ
2	5.31	4.66
4	5.27	4.60
6	5.81	4.87
8	5.86	5.18
10	7.96	6.67

Table C.7: E_{rel} values assessed across one-compartment Case 1 at each undersampling level using tkPINNs method.

Case	Dt	E_{rel}	
		u	
		μ	σ
1	2	4.28	4.41
1	4	4.27	4.37
1	6	5.48	5.02
1	8	5.00	4.86
1	10	11.74	8.38

Table C.8: E_{rel} values assessed across one-compartment Case 2 at each undersampling level using tkPINNs method.

Case	Dt	E_{rel}	
		u	
		μ	σ
2	2	6.68	5.46
2	4	6.75	5.43
2	6	6.70	5.50
2	8	6.71	5.58
2	10	7.03	4.89

Table C.9: E_{rel} values assessed across one-compartment Case 3 at each undersampling level using tkPINNs method.

Case	Dt	E_{rel}	
		u	
		μ	σ
3	2	4.97	3.58
3	4	4.81	3.41
3	6	5.25	3.81
3	8	5.87	4.93
3	10	5.09	4.01

Table C.10: E_{rel} values for each two-compartment systems case for noiseless data.

Case	E_{rel}					
	u^a		u^v		K^{va}	
	μ	σ	μ	σ	μ	σ
1	10.13	11.86	11.85	7.16	28.79	34.75
2	239.85	162.92	53.39	11.58	239.37	88.03
3	9.09	6.54	16.28	14.73	35.12	54.93
Overall	86.36	143.84	27.18	21.93	101.09	116.44

Table C.11: E_{rel} values assessed across all two-compartment system cases at each SNR level.

SNR	E_{rel}					
	u^a		u^v		K^{va}	
	μ	σ	μ	σ	μ	σ
5	95.89	52.76	94.37	49.59	92.21	105.91
10	65.28	49.43	86.02	45.01	73.58	105.86
15	58.71	91.96	80.54	42.29	99.74	114.91
20	1329.19	2410.84	80.80	44.55	1304.21	1840.66

Table C.12: E_{rel} values assessed across two-compartment Case 1 at each SNR level using tkPINNs method.

Case	SNR	E_{rel}					
		u^a		u^v		K^{va}	
		μ	σ	μ	σ	μ	σ
1	5	93.74	50.13	103.88	55.33	98.23	54.50
1	10	65.05	49.10	85.71	45.89	70.67	57.44
1	15	22.55	22.99	74.69	39.04	48.99	57.34
1	20	18.60	21.75	76.40	42.61	48.98	52.00

Table C.13: E_{rel} values assessed across two-compartment Case 2 at each SNR level using tkPINNs method.

Case	SNR	E_{rel}					
		u^a		u^v		K^{va}	
		μ	σ	μ	σ	μ	σ
2	5	99.94	50.69	98.98	44.58	98.93	166.88
2	10	97.30	47.64	97.25	40.94	100.36	165.97
2	15	142.31	119.09	93.73	40.19	201.35	133.23
2	20	3953.11	2666.15	94.14	42.41	3819.52	817.76

Table C.14: E_{rel} values assessed across two-compartment Case 3 at each SNR level using tkPINNs method.

Case	SNR	E_{rel}					
		u^a		u^v		K^{va}	
		μ	σ	μ	σ	μ	σ
3	5	94.00	56.97	80.27	44.91	79.48	50.86
3	10	33.50	24.79	75.10	45.30	49.72	38.43
3	15	11.28	10.65	73.20	44.31	48.89	55.57
3	20	15.87	11.92	71.86	45.40	44.13	49.69

Table C.15: E_{rel} values assessed across two-compartment Case 1 at each undersampling level using tkPINNs method.

Case	Dt	E_{rel}					
		u^a		u^v		K^{va}	
		μ	σ	μ	σ	μ	σ
1	2	10.13	11.86	11.85	7.16	28.79	34.75
1	4	8.35	10.85	7.31	4.31	19.23	23.02
1	6	14.35	17.89	36.57	15.78	43.85	40.31
1	8	19.21	20.38	61.15	32.56	52.10	42.43
1	10	29.96	22.70	75.75	47.15	49.56	35.57

Table C.16: E_{rel} values assessed across two-compartment Case 2 at each undersampling level using tkPINNs method.

Case	Dt	E_{rel}					
		u^a		u^v		K^{va}	
		μ	σ	μ	σ	μ	σ
2	2	239.85	162.92	53.39	11.58	239.37	88.03
2	4	344.13	225.67	50.60	10.10	313.04	116.65
2	6	460.36	302.33	44.65	7.52	406.99	159.95
2	8	358.04	241.90	40.36	7.36	352.75	131.93
2	10	160.87	116.22	70.34	23.22	177.42	63.87

Table C.17: E_{rel} values assessed across two-compartment Case 3 at each undersampling level using tkPINNs method.

Case	Dt	E_{rel}					
		u^a		u^v		K^{va}	
		μ	σ	μ	σ	μ	σ
3	2	9.09	6.54	16.28	14.73	35.12	54.93
3	4	17.91	15.69	50.29	24.41	53.51	56.07
3	6	15.35	10.72	40.81	15.35	77.84	122.04
3	8	21.74	18.03	109.44	79.32	143.63	162.84
3	10	18.71	17.71	22.06	15.69	61.25	61.91

Appendix D

Paper Abstracts

In the interest of thesis flow the abstract from the OSIPI DCE-Challenge paper (Shalom *et al.*, 2024a) is not included in Chapter 3. The structured abstract as accepted for publication is included below for completeness.

Abstract; The ISMRM Open Science Initiative for Perfusion Imaging (OSIPI): Results from the OSIPI-DCE Challenge.

Purpose: K^{trans} has often been proposed as a quantitative imaging biomarker for diagnosis, prognosis, and treatment response assessment for various tumors. None of the many software tools for K^{trans} quantification are standardized. The ISMRM OSIPI-DCE challenge was designed to benchmark methods to better help the efforts to standardize K^{trans} measurement.

Methods: A framework was created to evaluate K^{trans} values produced by DCE-MRI analysis pipelines to enable benchmarking. The perfusion MRI community was invited to apply their pipelines for K^{trans} quantification in glioblastoma from clinical and synthetic patients. Submissions were required to include the entrants' K^{trans} values, the applied software, and a standard operating procedure (SOP). These were evaluated using the proposed $OSIPI_{gold}$ score defined with accuracy, repeatability and reproducibility components.

Results: Across the 10 received submissions, the $OSIPI_{gold}$ score ranged from 28-78% with a 59% median. The accuracy, repeatability and reproducibility scores ranged from 0.54-0.92, 0.64-0.86 and 0.65-1.00, respectively (0-1=lowest-highest). Manual arterial input function (AIF)

selection markedly affected the reproducibility and showed greater variability in K^{trans} analysis than automated methods. Furthermore, provision of a detailed SOP was critical for higher reproducibility.

Conclusions: This study reports results from the OSIPI-DCE challenge and highlights the high inter-software variability within K^{trans} estimation, providing a framework for ongoing benchmarking against the scores presented. Through this challenge, the participating teams were ranked based on the performance of their software tools in the particular setting of this challenge. In a real-world clinical setting, many of these tools may perform differently with different benchmarking methodology.

Bibliography

Aarts, E. & Korst, J. (1989). *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing*. John Wiley & Sons, Inc., USA.

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y. & Zheng, X. (2015). TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. https://www.tensorflow.org/.

Al-Muhanna, A.F. (2022). Gadolinium Retention after Contrast-Enhanced Magnetic Resonance Imaging: A Narratative Review. *Saudi journal of medicine medical sciences*, **10**, 12–18, doi:10.4103/sjmms.sjmms_198_21.

Alsop, D.C., Detre, J.A., Golay, X., Günther, M., Hendrikse, J., Hernandez-Garcia, L., Lu, H., MacIntosh, B.J., Parkes, L.M., Smits, M., van Osch, M.J.P., Wang, D.J.J., Wong, E.C. & Zaharchuk, G. (2015). Recommended implementation of arterial spin-labeled perfusion MRI for clinical applications: A consensus of the ISMRM

perfusion study group and the European consortium for ASL in dementia. *Magnetic Resonance in Medicine*, **73**, 102–116, doi:10.1002/mrm.25197.

Banerji, A., Naish, J.H., Watson, Y., Jayson, G.C., Buonaccorsi, G.A. & Parker, G.J. (2012). DCE-MRI model selection for investigating disruption of microvascular function in livers with metastatic disease. *Journal of Magnetic Resonance Imaging*, **35**, 196–203, doi:10.1002/JMRI.22692.

Barboriak, D. (2015). Data From RIDER_NEURO_MRI. *The Cancer Imaging Archive*, doi:10.7937/K9/TCIA.2015.VOSN3HN1.

Barnes, S.L., Quarles, C.C. & Yankeelov, T.E. (2014). Modeling the effect of intra-voxel diffusion of contrast agent on the quantitative analysis of dynamic contrast enhanced magnetic resonance imaging. *PLoS ONE*, **9**, e108726, doi:10.1371/journal.pone.0108726.

Baydin, A., Pearlmutter, B., Radul, A. & Siskind, J. (2018). Automatic Differentiation in Machine Learning: a Survey. *Journal of Machine Learning Research*, **18**, 1–43.

Bell, L., Ragunathan, S. & Kazerooni, A.F. (2021). *Chapter 13 - Contrast agent-based perfusion MRI methods*, vol. 4 of *Advanced Neuro MR Techniques and Applications*, 195–209. Academic Press, doi:10.1016/B978-0-12-822479-3.00024-5.

Bergner, P.E., Lushbaugh, C.C. & Anderson, E.B. (1967). Compartments, pools, and spaces in medical physiology. AEC Symposium Series 11. Proceedings of a Symposium, October 24–27, 1966.

Berks, M. (2022). madym_cxx. *GitHub*, https://gitlab.com/manchester_qbi/manchester_qbi_public/osipi-dce-challenge.

Berks, M., Parker, G., Little, R. & Cheung, S. (2021). Madym: A C++ toolkit for quantitative DCE-MRI analysis. *Journal of Open Source Software*, **6**, 3523, doi:10.21105/joss.03523.

Beuzit, L., Eliat, P., Brun, V., Ferré, J., Gandon, Y., Bannier, E. & Saint-Jalmes, H. (2016). Dynamic contrast-enhanced MRI: Study of inter-software accuracy and reproducibility using simulated and clinical data. *Journal of Magnetic Resonance Imaging*, **43**, 1288–1300, doi:10.1002/jmri.25101.

Bhargava, V., Singh, K., Meena, P. & Sanyal, R. (2021). Nephrogenic systemic fibrosis: A frivolous entity. *World journal of nephrology*, **10**, 29–36, doi:10.5527/wjn.v10.i3.29.

Bloch, F., Hansen, W.W. & Packard, M. (1946). Nuclear induction. *Physical Review*, **69**, 127, doi:10.1103/PhysRev.69.127.

Blockley, N., Jiang, L., Gardener, A., Ludman, C., Francis, S. & Gowland, P. (2008). Field strength dependence of R_1 and R_2^* relaxivities of human whole blood to prohance, vasovist, and deoxyhemoglobin. *Magn Reson Med*, **60**, 1313–1320, doi:10.1002/mrm.21792.

Boucher, Y., Baxter, L.T. & Jain, R.K. (1990). Interstitial pressure gradients in tissue-isolated and subcutaneous tumors: implications for therapy. *Cancer research*, **50**, 4478–84, PMID: 2369726.

Boxerman, J.L., Quarles, C.C., Hu, L.S., Erickson, B.J., Gerstner, E.R., Smits, M., Kaufmann, T.J., Barboriak, D.P., Huang, R.H., Wick, W., Weller, M., Galanis, E., Kalpathy-Cramer, J., Shankar, L., Jacobs, P., Chung, C., van den Bent, M.J., Chang, S., Yung, W.K.A., Cloughesy, T.F., Wen, P.Y., Gilbert, M.R., Rosen, B.R., Ellingson, B.M., Schmainda, K.M., Arons, D.F., Kingston, A., Sandak, D., Wallace, M., Musella, A. & Haynes, C. (2020). Consensus recommendations for a dynamic susceptibility contrast MRI protocol for use in high-grade gliomas. *Neuro-Oncology*, **22**, 1262–1275, doi:10.1093/neuonc/noaa141.

Brett, M., Markiewicz, C., Hanke, M., Côté, M., Cipollini, B., McCarthy, P., Jarecka, D., Cheng, C., Halchenko, Y., Cottaar,

M., Larson, E., Ghosh, S., Wassermann, D., Gerhard, S., Lee, G., Wang, H., Kastman, E., Kaczmarzyk, J., Guidotti, R., Duek, O., Daniel, J., Rokem, A., Madison, C., Moloney, B., Morency, F.C., Goncalves, M., Markello, R., Riddell, C., Burns, C., Millman, J., Gramfort, A., Leppäkangas, J., Sólon, A., van den Bosch, J.J., Vincent, R.D., Braun, H., Subramaniam, K., Gorgolewski, K.J., Raamana, P.R., Klug, J., Nichols, B.N., Baker, E.M., Hayashi, S., Pinsard, B., Haselgrove, C., Hymers, M., Esteban, O., Koudoro, S., Pérez-García, F., Oosterhof, N.N., Amirbekian, B., Nimmo-Smith, I., Nguyen, L., Reddigari, S., St-Jean, S., Panfilov, E., Garyfallidis, E., Varoquaux, G., Legarreta, J.H., Hahn, K.S., Hinds, O.P., Fauber, B., Poline, J.B., Stutters, J., Jordan, K., Cieslak, M., Moreno, M.E., Haenel, V., Schwartz, Y., Baratz, Z., Darwin, B.C., Thirion, B., Gauthier, C., Orfanos, D.P., Solovey, I., Gonzalez, I., Palasubramaniam, J., Lecher, J., Leinweber, K., Raktivan, K., Calábková, M., Fischer, P., Gervais, P., Gadde, S., Ballinger, T., Roos, T. & Reddam, V. (2020). nipy/nibabel: 3.2.1. doi:10.5281/ZENODO.4295521.

Brix, G., Semmler, W., Port, R., Schad, L.R., Layer, G. & Lorenz, W.J. (1991). Pharmacokinetic Parameters in CNS Gd-DTPA Enhanced MR Imaging. *Journal of Computer Assisted Tomography*, **15**, 621–628, doi:10.1097/00004728-199107000-00018.

Brix, G., Kiessling, F., Lucht, R., Darai, S., Wasser, K., Delorme, S. & Griebel, J. (2004). Microcirculation and microvasculature in breast tumors: Pharmacokinetic analysis of dynamic MR image series. *Magn Reson Med*, **52**, 420–429, doi:10.1002/mrm.20161.

Brunton, S.L. & Kutz, J.N. (2019). *Neural Networks and Deep Learning*, 195–226. Cambridge University Press, doi:10.1017/9781009089517.009.

Brunton, S.L. & Kutz, J.N. (2022). *Data-Driven Science and Engineering*. Cambridge University Press, doi:10.1017/9781009089517.

Buckley, D.L. (2002). Uncertainty in the analysis of tracer kinetics using dynamic contrast-enhanced T_1-weighted MRI. *Magnetic Resonance in Medicine*, **47**, 601–606, doi:10.1002/mrm.10080.

Calamante, F. (2013). Arterial input function in perfusion MRI: A comprehensive review. *Progress in Nuclear Magnetic Resonance Spectroscopy*, doi:10.1016/j.pnmrs.2013.04.002.

Calamante, F., Gadian, D.G. & Connelly, A. (2000). Delay and dispersion effects in dynamic susceptibility contrast MRI: Simulations using singular value decomposition. *Magnetic Resonance in Medicine*, **44**, 466–473, doi:10.1002/1522-2594(200009)44:3<466::AID-MRM18>3.0.CO;2-M.

Calamante, F., Willats, L., Gadian, D.G. & Connelly, A. (2006). Bolus delay and dispersion in perfusion MRI: Implications for tissue predictor models in stroke. *Magnetic Resonance in Medicine*, **55**, 1180–1185, doi:10.1002/mrm.20873.

Carp, J. (2012). The secret lives of experiments: Methods reporting in the fMRI literature. *Neuroimage*, **63**, 289–300, doi:10.1016/j.neuroimage.2012.07.004.

Cheng, H.L.M. (2008). Investigation and optimization of parameter accuracy in dynamic contrast-enhanced MRI. *Journal of Magnetic Resonance Imaging*, **28**, 736–743, doi:10.1002/jmri.21489.

Chih-Feng, C., Ling-Wei, H., Chun-Chung, L., Chen-Chang, L., Hsu-Huei, W., Yuan-Hsiung, T. & Ho-Ling, L. (2012). In vivo correlation between semi-quantitative hemodynamic parameters and K trans derived from DCE-MRI of brain tumors. *International Journal of Imaging Systems and Technology*, **22**, 132–136, doi:10.1002/ima.22013.

Cho, H., Ackerstaff, E., Carlin, S., Lupu, M.E., Wang, Y., Rizwan, A., O'Donoghue, J., Ling, C.C., Humm, J.L., Zanzonico, P.B. & Koutcher, J.A. (2009). Noninvasive Multimodality Imaging of the Tumor Microenvironment: Registered Dynamic Magnetic Resonance Imaging and

Positron Emission Tomography Studies of a Preclinical Tumor Model of Tumor Hypoxia. *Neoplasia*, **11**, 247–IN3, doi:10.1593/neo.81360.

Christensen, S., Calamante, F., Hjort, N., Wu, O., Blankholm, A.D., Desmond, P., Davis, S. & Østergaard, L. (2008). Inferring origin of vascular supply from tracer arrival timing patterns using bolus tracking MRI. *Journal of Magnetic Resonance Imaging*, **27**, 1371–1381, doi:10.1002/jmri.21386.

Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., Tarbox, L. & Prior, F. (2013). The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *J Digit Imaging*, **26**, 1045–1057, doi:10.1007/s10278-013-9622-7.

Clement, P., Castellaro, M., Okell, T.W., Thomas, D.L., Vandemaele, P., Elgayar, S., Oliver-Taylor, A., Kirk, T., Woods, J.G., Vos, S.B., Kuijer, J.P.A., Achten, E., van Osch, M.J.P., Appelhoff, S., Blair, R., Feingold, F., Gau, R., Markiewicz, C.J., Salo, T., Detre, J.A., Lu, H., Alsop, D.C., Chappell, M.A., Hernandez-Garcia, L., Petr, J. & Mutsaerts, H.J.M.M. (2022). ASL-BIDS, the brain imaging data structure extension for arterial spin labeling. *Scientific Data*, **9**, 543, doi:10.1038/s41597-022-01615-9.

Cramer, S., Simonsen, H., Frederiksen, J., Rostrup, E. & Larsson, H. (2014). Abnormal blood–brain barrier permeability in normal appearing white matter in multiple sclerosis investigated by MRI. *NeuroImage Clin*, **4**, 182–189, doi:10.1016/j.nicl.2013.12.001.

Cuenod, C. & Balvay, D. (2013). Perfusion and vascular permeability: Basic concepts and measurement in DCE-CT and DCE-MRI. *Diagnostic and Interventional Imaging*, **94**, 1187–1204, doi:https://doi.org/10.1016/j.diii.2013.10.010.

Cuomo, S., Cola, V.S.D., Giampaolo, F., Rozza, G., Raissi, M. & Piccialli, F. (2022). Scientific Machine Learning Through Physics–Informed Neural Networks: Where we are and What's Next. *Journal of Scientific Computing*, **92**, 88, doi:10.1007/s10915-022-01939-z.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*, **2**, 303–314, doi:10.1007/BF02551274.

Damadian, R. (1971). Tumor detection by nuclear magnetic resonance. *Science*, **171**, 1151–1153, doi:10.1126/science.171.3976.1151.

Damadian, R., Goldsmith, M. & Minkoff, L. (1977). NMR in cancer: XVI. FONAR image of the live human body. *Physiological chemistry and physics*, **9**, 97–100, 108.

Demeestere, J., Wouters, A., Christensen, S., Lemmens, R. & Lansberg, M.G. (2020). Review of Perfusion Imaging in Acute Ischemic Stroke. *Stroke*, **51**, 1017–1024, doi:10.1161/STROKEAHA.119.028337.

Dickie, B.R., Ahmed, Z., Arvidsson, J., Bell, L.C., Buckley, D.L., Debus, C., Fedorov, A., Floca, R., Gutmann, I., van der Heijden, R.A., van Houdt, P.J., Sourbron, S., Thrippleton, M.J., Quarles, C. & Kompan, I.N. (2023). A community-endorsed open-source lexicon for contrast agent–based perfusion MRI : A consensus guidelines report from the ISMRM Open Science Initiative for Perfusion Imaging (OSIPI). *Magnetic Resonance in Medicine*, doi:10.1002/mrm.29840.

Don, T.D.J., Safaei, S., Talou, G.D.M., Russell, P.S., Phillips, A.R.J. & Reynolds, H.M. (2024). Computational fluid dynamic modeling of the lymphatic system: a review of existing models and future directions. *Biomechanics and Modeling in Mechanobiology*, **23**, 3–22, doi:10.1007/s10237-023-01780-9.

Egeland, T.A., Gulliksrud, K., Gaustad, J.V., Mathiesen, B. & Rofstad, E.K. (2012). Dynamic contrast-enhanced-MRI of tumor hypoxia. *Magnetic Resonance in Medicine*, **67**, 519–530, doi:10.1002/mrm.23014.

Egeland, T.A.M., Gaustad, J.V., Benjaminsen, I.C., Hedalen, K., Mathiesen, B. & Rofstad, E.K. (2008). Assessment of Fraction of Hypoxic Cells in Human Tumor Xenografts with Necrotic Regions by Dynamic Contrast-Enhanced MRI. *Radiation Research*, **169**, 689–699, doi:10.1667/RR1311.1.

Egeland, T.A.M., Gaustad, J.V., Galappathi, K. & Rofstad, E.K. (2011). Magnetic resonance imaging of tumor necrosis. *Acta Oncologica*, **50**, 427–434, doi:10.3109/0284186X.2010.526633.

Elkin, R., Nadeem, S., LoCastro, E., Paudyal, R., Hatzoglou, V., Lee, N.Y., Shukla-Dave, A., Deasy, J.O., Tannenbaum, A. & equally contributing authors, S.N. (2019). Optimal mass transport kinetic modeling for head and neck DCE-MRI: Initial analysis. *Magn Reson Med*, **82**, 2314–2325, doi:10.1002/mrm.27897.

Elsharif, M., Roche, M., Wilson, D., Basak, S., Rowe, I., Vijayanand, D., Feltbower, R., Treanor, D., Roberts, L., Guthrie, A., Prasad, R., Gilthorpe, M.S., Attia, M. & Sourbron, S. (2021). Hepatectomy risk assessment with functional magnetic resonance imaging (heparim). *BMC Cancer*, **21**, 1139, doi:10.1186/s12885-021-08830-4.

FDA-NIH Biomarker Working Group (2016). *BEST (Biomarkers, EndpointS, and other Tools) Resource*. Silver Spring (MD): Food and Drug Administration (US). Co-published by National Institutes of Health (US), Bethesda (MD), www.ncbi.nlm.nih.gov/books/NBK326791/, Accessed 09.02.2024.

Flouri, D. (2016). *Tracer-Kinetic Model-Driven Motion Correction with Application to Renal DCE-MRI*. Ph.D. thesis.

Flouri, D., Lesnic, D. & Sourbron, S.P. (2016). Fitting the two-compartment model in DCE-MRI by linear inversion. *Magnetic Resonance in Medicine*, **76**, 998–1006, doi:10.1002/mrm.25991.

Fluckiger, J.U., Schabel, M.C. & DiBella, E.V. (2009). Model-based blind estimation of kinetic parameters in dynamic contrast enhanced (dce)-mri. *Magnetic Resonance in Medicine*, **62**, 1477–1486, doi:10.1002/mrm.22101.

Fluckiger, J.U., Loveless, M.E., Barnes, S.L., Lepage, M., Yankeelov, T.E. & Edu, J.F. (2013). A diffusion-compensated model for the analysis of DCE-MRI data: theory, simulations, and experimental results. *Phys Med Biol*, **58**, 1983–1998, doi:10.1088/0031-9155/58/6/1983.

Fram, E.K., Herfkens, R.J., Johnson, G., Glover, G.H., Karis, J.P., Shimakawa, A., Perkins, T.G. & Pelc, N.J. (1987). Rapid calculation of t1 using variable flip angle gradient refocused imaging. *Magnetic Resonance Imaging*, **5**, 201–208, doi:10.1016/0730-725X(87)90021-X.

Frank, L.R., Lu, K. & Wong, E.C. (2008). Perfusion tensor imaging. *Magnetic Resonance in Medicine*, **60**, 1284–1291, doi:10.1002/mrm.21806.

Gejadze, I.Y., Dimet, F.X.L. & Shutyaev, V. (2008). On analysis error covariances in variational data assimilation. *SIAM Journal on Scientific Computing*, **30**, 1847–1874, doi:10.1137/07068744X.

Griewank, A. & Walther, A. (2008). In *Evaluating Derivatives*, Other Titles in Applied Mathematics, Society for Industrial and Applied Mathematics, doi:10.1137/1.9780898717761.

Gurney-Champion, O.J. & Ottens, T. (2022). DCE-NET. *GitHub*, https://github.com/oliverchampion/DCENET.

Hanson, E.A., Sandmann, C., Malyshev, A., Lundervold, A., Modersitzki, J. & Hodneland, E. (2018). Estimating the discretization dependent accuracy of perfusion in coupled capillary flow measurements. *PLOS ONE*, **13**, e0200521, doi:10.1371/journal.pone.0200521.

Hanson, L.G. (2008). Is quantum mechanics necessary for understanding magnetic resonance? *Concepts in Magnetic Resonance Part A*, **32A**, 329–340, doi:10.1002/cmr.a.20123.

Hanson, L.G. (2015). *The Ups and Downs of Classical and Quantum Formulations of Magnetic Resonance*, 141–171. Anthropic Awareness: The Human Aspects of Scientific Thinking in NMR Spectroscopy and Mass Spectrometry, Elsevier Inc., doi:10.1016/B978-0-12-419963-7.00003-1.

Harris, C., Millman, K., van der Walt, S., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M., Brett, M., Haldane, A., del Río, J., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C. & Oliphant, T. (2020). Array programming with numpy. *Nature*, **585**, 357–362, doi:10.1038/s41586-020-2649-2.

Henkelman, R.M. (1990). Does IVIM measure classical perfusion? *Magnetic Resonance in Medicine*, **16**, 470–475, doi:10.1002/mrm.1910160313.

Hernandez-Garcia, L., Lahiri, A. & Schollenberger, J. (2019). Recent progress in ASL. *NeuroImage*, **187**, 3–16, doi:10.1016/j.neuroimage.2017.12.095.

Hunter-Zinck, H., de Siqueira, A.F., Vásquez, V.N., Barnes, R. & Martinez, C.C. (2021). Ten simple rules on writing clean and reliable open-source scientific software. *PLOS Computational Biology*, **17**, e1009481, doi:10.1371/journal.pcbi.1009481.

Iglesias, M.A., Law, K.J. & Stuart, A.M. (2013). Ensemble Kalman Methods for Inverse Problems. Tech. rep., University of Warwick.

Ingrisch, M. & Sourbron, S. (2013). Tracer-kinetic modeling of dynamic contrast-enhanced MRI and CT: A primer. *Journal of Pharmacokinetics and Pharmacodynamics*, **40**, 281–300, doi:10.1007/s10928-013-9315-3.

Jagtap, A.D., Kharazmi, E. & Karniadakis, G.E. (2020). Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems. *Computer Methods in Applied Mechanics and Engineering*, **365**, 113028, doi:10.1016/j.cma.2020.113028.

Jahng, G.H., Li, K.L., Ostergaard, L. & Calamante, F. (2014). Perfusion Magnetic Resonance Imaging: A Comprehensive Update on Principles and Techniques. *Korean J Radiol*, **15**, 554–577, doi:10.3348/kjr.2014.15.5.554.

Jiřík, R., Taxt, T., Macíček, O., Bartoš, M., Kratochvíla, J., Souček, K., Dražanová, E., Krátká, L., Hampl, A. & Starčuk, Z. (2019). Blind deconvolution estimation of an arterial input function for small animal DCE-MRI. *Magn Reson Imaging*, **62**, 46–56, doi:10.1016/j.mri.2019.05.024.

Jiřík, R., Bartoš, M., Macíček, O. & Starčuk, Z. (2022). Arterial Input Function Estimation Using All-Channel Blind Deconvolution with Spatial Regularization in DCE-MRI. *Proc. Intl. Soc. Mag. Reson. Med.*, 1582–1582.

Jiřík, R. (2022). Perflab. *CERIT Scientific Cloud*, http://perflab.cerit-sc.cz/.

Kajishima, T., Taira, K., Kajishima, T. & Taira, K. (2017). *Finite-Difference Discretization of the Advection-Diffusion Equation*, 23–72. Springer International Publishing, doi:10.1007/978-3-319-45304-0_2.

Kazerooni, A., Bell, L., den Abeele, F.V., Verhack, R., Zhou, X., Rezaei, S., Ahmed, Z., Heijden, R.V.D., Nabavizadeh, S., Hu, L., Rad, H. & Sourbron, S. (2021a). The Open Source Initiative for Perfusion Imaging (OSIPI): DCE-MRI Challenge. *Proc Intl Soc Mag Reson Med*, 1094–1094.

Kazerooni, A., Shalom, E., Ahmed, Z., van der Heijden, R., Zhao, M., Kim, H., Bell, L., Sourbron, S., Laue, H. & Schouten, P. (2021b). OSIPI DCE Challenge. https://osf.io/u7a6f/, doi:10.17605/OSF.IO/U7A6F.

Kety, S. (1951). The theory and applications of the exchange of inert gas at the lungs and tissues. *Pharmacological reviews*, **3**, 1–41.

Khalifa, F., Soliman, A., El-Baz, A., El-Ghar, M.A., El-Diasty, T., Gimel'farb, G., Ouseph, R. & Dwyer, A.C. (2014). Models and methods for analyzing DCE-MRI: A review. *Medical Physics*, **41**, 124301, doi:10.1118/1.4898202.

Kickingereder, P., Wiestler, B., Graf, M., Heiland, S., Schlemmer, H., Wick, W., Wick, A., Bendszus, M. & Radbruch, A. (2015). Evaluation of dynamic contrast-enhanced MRI derived microvascular permeability in recurrent glioblastoma treated with bevacizumab. *J Neurooncol*, **121**, 373–380, doi:10.1007/s11060-014-1644-6.

Kim, H. (2018). Variability in Quantitative DCE-MRI: Sources and Solutions. *Journal of nature and science*, **4**.

Kingma, D.P. & Ba, J. (2015). Adam: A Method for Stochastic Optimization. *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.

Kirkpatrick, S., Gelatt, C.D. & Vecchi, M.P. (1983). Optimization by Simulated Annealing. *Science*, **220**, 671–680, doi:10.1126/science.220.4598.671.

Kissas, G., Yang, Y., Hwuang, E., Witschey, W.R., Detre, J.A. & Perdikaris, P. (2020). Machine learning in cardiovascular flows modeling: Predicting arterial blood pressure from non-invasive 4D flow MRI data using physics-informed neural networks. *Computer Methods in Applied Mechanics and Engineering*, **358**, 112623, doi:10.1016/j.cma.2019.112623.

Klawer, E.M., van Houdt, P.J., Pos, F.J., Heijmink, S.W., van Osch, M.J. & van der Heide, U.A. (2018). Impact of contrast agent injection duration on dynamic contrast-enhanced MRI quantification in prostate cancer. *NMR in Biomedicine*, **31**, e3946, doi:10.1002/nbm.3946.

Koeppe, A., Bamer, F. & Markert, B. (2020). An intelligent nonlinear meta element for elastoplastic continua: deep learning using a new Time-distributed Residual U-Net architecture. *Computer Methods in Applied Mechanics and Engineering*, **366**, 113088, doi:10.1016/j.cma.2020.113088.

Koh, T.S., Bisdas, S., Koh, D.M. & Thng, C.H. (2011). Fundamentals of tracer kinetics for dynamic contrast-enhanced MRI. *Journal of Magnetic Resonance Imaging*, **34**, 1262–1276, doi:10.1002/jmri.22795.

Koh, T.S., Hartono, S., Thng, C.H., Lim, T.K., Martarello, L. & Ng, Q.S. (2013). In vivo measurement of gadolinium diffusivity by dynamic contrast-enhanced MRI: A preclinical study of human xenografts. *Magnetic Resonance in Medicine*, **69**, 269–276, doi:10.1002/mrm.24246.

Krishnapriyan, A.S., Gholami, A., Zhe, S., Kirby, R.M. & Mahoney, M.W. (2021). Characterizing possible failure modes in physics-informed neural networks. *35th Conference on Neural Information Processing Systems (NeurIPS 2021)*, doi:10.48550/arXiv.2109.01050.

Kuhl, C.K., Mielcareck, P., Klaschik, S., Leutner, C., Wardelmann, E., Gieseke, J. & Schild, H.H. (1999). Dynamic Breast MR Imaging: Are Signal Intensity Time Course Data Useful for Differential Diagnosis of Enhancing Lesions? *Radiology*, **211**, 101–110, pMID: 10189459, doi:10.1148/radiology.211.1.r99ap38101.

Larsson, H., Courivaud, F., Rostrup, E. & Hansen, A. (2009). Measurement of brain perfusion, blood volume, and blood-brain barrier permeability, using dynamic contrast-enhanced T_1 weighted MRI at 3 tesla. *Magn Reson Med*, **62**, 1270–1281, doi:10.1002/mrm.22136.

Larsson, H.B.W., Stubgaard, M., Frederiksen, J.L., Jensen, M., Henriksen, O. & Paulson, O.B. (1990). Quantitation of blood-brain barrier defect by magnetic resonance imaging and gadolinium-dtpa in patients with multiple sclerosis and brain tumors. *Magnetic Resonance in Medicine*, **16**, 117–131, doi:10.1002/mrm.1910160111.

Lassen, N.A. & Perl, W. (1979). *Tracer kinetic methods in medical physiology*. Raven Press.

Lauterbur, P.C. (1973). Image formation by induced local interactions: Examples employing nuclear magnetic resonance. *Nature*, **242**, 190–191, doi:10.1038/242190a0.

Lauterbur, P.C., Dias, M., Mendonça, H. & Rudin, A.M. (1978). *Augmentation of Tissue Water Proton Spin-lattice Relaxation Rates by In Vivo Addition of Paramagnetic Ions*, 752–759. Frontiers of Biological Energetics, Volume I: Electrons to Tissues, Elsevier, doi:10.1016/b978-0-12-225401-7.50093-x.

Leiteritz, R. & Pflüger, D. (2021). How to Avoid Trivial Solutions in Physics-Informed Neural Networks. *arXiv CoRR*, doi:10.48550/arXiv.2112.05620.

Levitt, M.H. (2013). *Spin Dynamics: Basics of Nuclear Magnetic Resonance*. Wiley, 2nd edn.

Lindquist, M. (2020). Neuroimaging results altered by varying analysis pipelines. *Nature*, **582**, 36–37, doi:10.1038/d41586-020-01282-z.

Liu, P., Lee, Y.Z., Aylward, S.R. & Niethammer, M. (2021a). Perfusion Imaging: An Advection Diffusion Approach. *IEEE Transactions on Medical Imaging*, **40**, 3424–3435, doi:10.1109/TMI.2021.3085828.

Liu, P., Tian, L., Zhang, Y., Aylward, S., Lee, Y. & Niethammer, M. (2021b). Discovering Hidden Physics Behind Transport Dynamics. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 10077–10087, doi:10.1109/CVPR46437.2021.00995.

Liu, P., Lee, Y., Aylward, S. & Niethammer, M. (2022). Deep Decomposition for Stochastic Normal-Abnormal Transport. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 18769–18779, doi:10.1109/CVPR52688.2022.01823.

Lohrke, J., Frenzel, T., Endrikat, J., Alves, F.C., Grist, T.M., Law, M., Lee, J.M., Leiner, T., Li, K.C., Nikolaou, K., Prince, M.R., Schild, H.H., Weinreb, J.C., Yoshikawa, K. & Pietsch, H. (2016). 25 Years of Contrast-Enhanced MRI: Developments, Current Challenges and Future Perspectives. *Advances in Therapy*, **33**, 1–28, doi:10.1007/s12325-015-0275-4.

Maes, F., Vandermeulen, D. & Suetens, P. (1999). Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Medical Image Analysis*, **3**, 373–386, doi:10.1016/S1361-8415(99)80030-9.

Mansfield, P. & Maudsley, A.A. (1977). Medical imaging by NMR. *British Journal of Radiology*, **50**, 188–194, doi:10.1259/0007-1285-50-591-188.

Mansfield, P. & Morris, P. (1982). *NMR Imaging in Biomedicine: Supplement 2 Advances in Magnetic Resonance*. Academic Press, 1st edn.

Matejka, J. & Fitzmaurice, G. (2017). Same stats, different graphs. *Proc SIGCHI Conf Hum Factor Comput Syst*, 1290–1294, doi:10.1145/3025453.3025912.

McCulloch, W.S. & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, **5**, 115–133, doi:10.1007/BF02478259.

McKiernan, E.C., Bourne, P.E., Brown, C.T., Buck, S., Kenall, A., Lin, J., McDougall, D., Nosek, B.A., Ram, K., Soderberg, C.K., Spies, J.R., Thaney, K., Updegrove, A., Woo, K.H. & Yarkoni, T. (2016). How open science helps researchers succeed. *eLife*, **5**, doi:10.7554/eLife.16800.

Mehlig, B. (2021). *Machine Learning with Neural Networks*. Cambridge University Press, doi:10.1017/9781108860604.

Mikheev, A. & Rusinek, H. (2022). FireVoxel. https://firevoxel.org/.

Milosevic, M., Fyles, A., Hedley, D., Pintilie, M., Levin, W., Manchul, L. & Hill, R. (2001). Interstitial fluid pressure predicts survival in patients with cervix cancer independent of clinical prognostic factors and tumor oxygen measurements. *Cancer research*, **61**, 6400–5, PMID: 11522633.

Milosevic, M.F., Pintilie, M., Hedley, D.W., Bristow, R.G., Wouters, B.G., Oza, A.M., Laframboise, S., Hill, R.P. & Fyles, A.W. (2014). High tumor interstitial fluid pressure identifies cervical cancer patients with improved survival from radiotherapy plus cisplatin versus radiotherapy alone. *International Journal of Cancer*, **135**, 1692–1699, doi:10.1002/ijc.28403.

Moore, E.H. (1920). On the reciprocal of the general algebraic matrix. *Bulletin of the American Mathematical Society*, **26**, 385–397, doi:10.1090/S0002-9904-1920-03322-7.

Murase, K. (2004). Efficient method for calculating kinetic parameters using T1-weighted dynamic contrast-enhanced magnetic resonance imaging. *Magn Reson Med*, **51**, 858–862, doi:10.1002/mrm.20022.

Naevdal, G., Saevareid, O. & Lorentzen, R.J. (2016). Data assimilation using mri data. *VII European Congress on Computational Methods in Applied Sciences and Engineering*.

Nair, V. & Hinton, G. (2010). Rectified Linear Units Improve Restricted Boltzmann Machines. *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*.

O'Connor, J., Jackson, A., Parker, G. & Jayson, G. (2007). DCE-MRI biomarkers in the clinical evaluation of antiangiogenic and vascular disrupting agents. *Br J Cancer*, **96**, 189–195, doi:10.1038/sj.bjc.6603515.

O'Connor, J.P.B., Aboagye, E.O., Adams, J.E., Aerts, H.J.W.L., Barrington, S.F., Beer, A.J., Boellaard, R., Bohndiek, S.E., Brady, M., Brown, G., Buckley, D.L., Chenevert, T.L., Clarke, L.P., Collette, S., Cook, G.J., deSouza, N.M., Dickson, J.C., Dive, C., Evelhoch, J.L., Faivre-Finn, C., Gallagher, F.A., Gilbert, F.J.,

Gillies, R.J., Goh, V., Griffiths, J.R., Groves, A.M., Halligan, S., Harris, A.L., Hawkes, D.J., Hoekstra, O.S., Huang, E.P., Hutton, B.F., Jackson, E.F., Jayson, G.C., Jones, A., Koh, D.M., Lacombe, D., Lambin, P., Lassau, N., Leach, M.O., Lee, T.Y., Leen, E.L., Lewis, J.S., Liu, Y., Lythgoe, M.F., Manoharan, P., Maxwell, R.J., Miles, K.A., Morgan, B., Morris, S., Ng, T., Padhani, A.R., Parker, G.J.M., Partridge, M., Pathak, A.P., Peet, A.C., Punwani, S., Reynolds, A.R., Robinson, S.P., Shankar, L.K., Sharma, R.A., Soloviev, D., Stroobants, S., Sullivan, D.C., Taylor, S.A., Tofts, P.S., Tozer, G.M., van Herk, M., Walker-Samuel, S., Wason, J., Williams, K.J., Workman, P., Yankeelov, T.E., Brindle, K.M., McShane, L.M., Jackson, A. & Waterton, J.C. (2017). Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol*, **14**, 169–186, doi:10.1038/nrclinonc.2016.162.

OSIPI (2023). Open Science Initiative for Perfusion Imaging (OSIPI). *GitHub*, https://github.com/OSIPI.

OSIPI Taskforce 4.1 (2023). Arterial Spin Labeling perfusion imaging and analysis lexicon and reporting recommendations. *Google Docs*, https://docs.google.com/document/d/e/2PACX-1vQPhnYciWScZ2vqjzBWUO72gXKKD2UA95q081OA7-2dnLiy9kOimXq1axA7E7xRxbZOmJAb9vMDZ_-D/pub.

Ottens, T., Barbieri, S., Orton, M.R., Klaassen, R., van Laarhoven, H.W., Crezee, H., Nederveen, A.J., Zhen, X. & Gurney-Champion, O.J. (2022). Deep learning DCE-MRI parameter estimation: Application in pancreatic cancer. *Medical Image Analysis*, **80**, 102512, doi:10.1016/j.media.2022.102512.

Parker, G., Roberts, C., Macdonald, A., Buonaccorsi, G., Cheung, S., Buckley, D., Jackson, A., Watson, Y., Davies, K. & Jayson, G. (2006). Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced mri. *Magn Reson Med*, **56**, 993–1000, doi:10.1002/mrm.21066.

Pautot, F. (2013). Theoretical Basis of in Vivo Tomographic Tracer Kinetics. http://viXra.org/abs/1304.0162.

Pellerin, M., Yankeelov, T.E. & Lepage, M. (2007). Incorporating contrast agent diffusion into the analysis of DCE-MRI data. *Magnetic Resonance in Medicine*, **58**, 1124–1134, doi:10.1002/mrm.21400.

Penrose, R. (1955). A generalized inverse for matrices. *Mathematical Proceedings of the Cambridge Philosophical Society*, **51**, 406–413, doi:10.1017/S0305004100030401.

Perl, W., Lassen, N.A. & Effros, R.M. (1975). Matrix proof of flow, volume and mean transit time theorems for regional and compartmental systems. *Bulletin of Mathematical Biology*, **37**, 573–588, doi:10.1007/BF02459526.

Petralia, G., Summers, P.E., Agostini, A., Ambrosini, R., Cianci, R., Cristel, G., Calistri, L. & Colagrande, S. (2020). Dynamic contrast-enhanced MRI in oncology: how we do it. *La radiologia medica*, **125**, 1288–1300, doi:10.1007/s11547-020-01220-z.

Peyrounette, M., Davit, Y., Quintard, M. & Lorthois, S. (2018). Multiscale modelling of blood flow in cerebral microcirculation: Details at capillary scale control accuracy at the level of the cortex. *PLOS ONE*, **13**, e0189474, doi:10.1371/journal.pone.0189474.

Pickens, D. (2000). *Magnetic Resonance Imaging*, 373–461. SPIE, doi:10.1117/3.832716.ch6.

Pintaske, J., Martirosian, P., Graf, H., Erb, G., Lodemann, K., Claussen, C. & Schick, F. (2006). Relaxivity of Gadopentetate Dimeglumine (Magnevist), Gadobutrol (Gadovist), and Gadobenate Dimeglumine (MultiHance) in Human Blood Plasma at 0.2, 1.5, and 3 Tesla. *Invest Radiol*, **41**, 213–221, doi:10.1097/01.rli.0000197668.44926.f7.

Port, R.E., Knopp, M.V. & Brix, G. (2001). Dynamic contrast-enhanced MRI using Gd-DTPA: Interindividual variability of the arterial input function

and consequences for the assessment of kinetics in tumors. *Magn Reson Med*, **45**, 1030–1038, doi:10.1002/mrm.1137.

Purcell, E.M., Torrey, H.C. & Pound, R.V. (1946). Resonance absorption by nuclear magnetic moments in a solid. *Physical Review*, **69**, 37–38, doi:10.1103/PhysRev.69.37.

QIBA MR Biomarker Committee (2020). MR DCE-MRI Quantification (DCEMRI-Q), Quantitative Imaging Biomarkers Alliance. Profile Stage: Public Comment. http://qibawiki.rsna.org/index.php/Profiles.

Raissi, M. & Karniadakis, G.E. (2018). Hidden physics models: Machine learning of nonlinear partial differential equations. *Journal of Computational Physics*, **357**, 125–141, doi:10.1016/j.jcp.2017.11.039.

Raissi, M., Perdikaris, P. & Karniadakis, G.E. (2017a). Physics Informed Deep Learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations. doi:10.48550/arXiv.1711.10561.

Raissi, M., Perdikaris, P. & Karniadakis, G.E. (2017b). Physics Informed Deep Learning (Part II): Data-driven Discovery of Nonlinear Partial Differential Equations. doi:10.48550/arXiv.1711.10566.

Raissi, M., Yazdani, A. & Karniadakis, G.E. (2018). Hidden Fluid Mechanics: A Navier-Stokes Informed Deep Learning Framework for Assimilating Flow Visualization Data. doi:10.48550/arXiv.1808.04327.

Raissi, M., Perdikaris, P. & Karniadakis, G.E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, **378**, 686–707, doi:10.1016/j.jcp.2018.10.045.

Raissi, M., Yazdani, A. & Karniadakis, G.E. (2020). Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science*, **367**, 1026–1030, doi:10.1126/science.aaw4741.

Rall, L.B. (1981). *Automatic Differentiation: Techniques and Applications*, vol. 120. Springer Berlin Heidelberg, doi:10.1007/3-540-10861-0.

Ramachandran, P., Zoph, B. & Le, Q.V. (2017). Searching for Activation Functions, https://arxiv.org/abs/1710.05941.

Rata, M., Collins, D.J., Darcy, J., Messiou, C., Tunariu, N., Desouza, N., Young, H., Leach, M.O. & Orton, M.R. (2016). Assessment of repeatability and treatment response in early phase clinical trials using DCE-MRI: comparison of parametric analysis using MR- and CT-derived arterial input functions. *European Radiology*, **26**, 1991–1998, doi:10.1007/s00330-015-4012-9.

Ronneberger, O., Fischer, P. & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, **9351**, 234–241, doi:10.1007/978-3-319-24574-4$_2$8.

Sainz-DeMena, D., Ye, W., Ángeles Pérez, M. & García-Aznar, J.M. (2022). A finite element based optimization algorithm to include diffusion into the analysis of DCE-MRI. *Engineering with Computers*, doi:10.1007/s00366-022-01667-w.

Schieda, N., Krishna, S. & Davenport, M.S. (2019). Update on Gadolinium-Based Contrast Agent–Enhanced Imaging in the Genitourinary System. *American Journal of Roentgenology*, **212**, 1223–1233, doi:10.2214/AJR.19.21137.

Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, **9**, 671–675, doi:10.1038/nmeth.2089.

Shah, A.H. & Olivero, J.J. (2017). Gadolinium-Induced Nephrogenic Systemic Fibrosis. *Methodist DeBakey cardiovascular journal*, **13**, 172–173, doi:10.14797/mdcj-13-3-172.

Shalom, E.S. (2023). tkspace, version 0.0.1. *GitHub*, https://github.com/EShalom/tkspace, doi:10.5281/zenodo.10056112.

Shalom, E.S., Khan, A., Van Loo, S. & Sourbron, S.P. (2024a). Current status in spatiotemporal analysis of contrast-based perfusion MRI. *Magnetic Resonance in Medicine*, **91**, 1136–1148, doi:10.1002/mrm.29906.

Shalom, E.S., Kim, H., van der Hijden, R., Ahmed, Z., Patel, R., II, D.A.H., DiCarlo, J.C., Yankeelov, T.E., Sisco, N.J., Dortch, R.D., Stokes, A.M., Inglese, M., Grech-Sollars, M., Toschi, N., Sahoo, P., Singh, A., Verma, S.K., Rathore, D.K., Kazerouni, A.S., Partridge, S.C., LoCastro, E., Paudyal, R., Wolansky, I.A., Shukla-Dave, A., Schouten, P., Gurney-Champion, O.J., Jiřík, R., Macíček, O., Bartoš, M., Vitouš, J., Das, A.B., Kim, S.G., Bokacheva, L., Mikheev, A., Rusinek, H., Berks, M., Cristinacce, P.L.H., Little, R.A., Cheung, S., O'Connor, J.P.B., Parker, G.J.M., Moloney, B., LaViolette, P.S., Bobholz, S., Duenweg, S., Virostko, J., Laue, H.O., Sung, K., Nabavizadeh, A., Rad, H.S., Hu, L.S., Sourbron, S.P., Bell, L.C. & Kazerooni, A.F. (2024b). The ISMRM Open Science Initiative for Perfusion Imaging (OSIPI): Results from the OSIPI-DCE Challenge. *Magnetic Resonance in Medicine*, **91**, 1803–1821, doi:10.1002/mrm.29909.

Shalom, E.S., Van Loo, S., Khan, A. & Sourbron, S.P. (2024c). Identifiability of spatiotemporal tissue perfusion models. *Physics in Medicine and Biology*, doi:10.1088/1361-6560/ad4087.

Sharma, P., Chung, W.T., Akoush, B. & Ihme, M. (2023). A Review of Physics-Informed Machine Learning in Fluid Mechanics. *Energies*, **16**, 2343, doi:10.3390/en16052343.

Shiroishi, M.S., Castellazzi, G., Boxerman, J.L., D'Amore, F., Essig, M., Nguyen, T.B., Provenzale, J.M., Enterline, D.S., Anzalone, N., Dörfler, A., Àlex Rovira, Wintermark, M. & Law, M. (2015). Principles of t_2^*-weighted dynamic susceptibility contrast MRI technique in brain tumor imaging. *Journal of Magnetic Resonance Imaging*, **41**, 296–313, doi:10.1002/jmri.24648.

Shukla-Dave, A., Obuchowski, N.A., Chenevert, T.L., Jambawalikar, S., Schwartz, L.H., Malyarenko, D., Huang, W., Noworolski, S.M., Young, R.J., Shiroishi, M.S., Kim, H., Coolens, C., Laue, H., Chung, C., Rosen, M., Boss, M. & Jackson, E.F. (2019). Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials. *J Magn Reson Imaging*, **49**, doi:10.1002/jmri.26518.

Siemonsen, S., Finsterbusch, J., Matschke, J., Lorenzen, A., Ding, X. & Fiehler, J. (2008). Age-Dependent Normal Values of T2* and T2 in Brain Parenchyma. *Am J Neuroradiol*, **29**, 950–955, doi:10.3174/ajnr.A0951.

Simeth, J., Johansson, A., Owen, D., Cuneo, K., Mierzwa, M., Feng, M., Lawrence, T.S. & Cao, Y. (2018). Quantification of liver function by linearization of a two-compartment model of gadoxetic acid uptake using dynamic contrast-enhanced magnetic resonance imaging. *NMR in Biomedicine*, **31**, e3913, doi:10.1002/nbm.3913.

Simmons, J., Nelson, L. & Simonsohn, U. (2011). False-positive psychology. *Psychol Sci*, **22**, 1359–1366, doi:10.1177/0956797611417632.

Singh, A., Rathore, R.K.S., Haris, M., Verma, S.K., Husain, N. & Gupta, R.K. (2009). Improved bolus arrival time and arterial input function estimation for tracer kinetic analysis in DCE-MRI. *Journal of Magnetic Resonance Imaging*, **29**, 166–176, doi:10.1002/jmri.21624.

Sinno, N., Taylor, E., Milosevic, M., Jaffray, D.A. & Coolens, C. (2021). Incorporating cross-voxel exchange into the analysis of dynamic contrast-enhanced imaging data: theory, simulations and experimental results. *Physics in Medicine Biology*, **66**, 205018, doi:10.1088/1361-6560/AC2205.

Sinno, N., Taylor, E., Hompland, T., Milosevic, M., Jaffray, D.A. & Coolens, C. (2022). Incorporating cross-voxel exchange for the analysis of dynamic contrast-enhanced imaging data: pre-clinical results. *Physics in Medicine Biology*, **67**, 245013, doi:10.1088/1361-6560/aca512.

Snyman, J.A. (2005). *Practical Mathematical Optimization*, vol. 97. Springer-Verlag, doi:10.1007/b105200.

Sourbron, S. (2014). A tracer-kinetic field theory for medical imaging. *IEEE Transactions on Medical Imaging*, **33**, 935–946, doi:10.1109/TMI.2014.2300450.

Sourbron, S. (2015). Reconstructing the one-compartment tracer-kinetic field with diffusion and convection. *Proc. Intl. Soc. Mag. Reson. Med*, **23**, 2324.

Sourbron, S.P. & Buckley, D.L. (2011). On the scope and interpretation of the Tofts models for DCE-MRI. *Magnetic Resonance in Medicine*, **66**, 735–745, doi:10.1002/mrm.22861.

Sourbron, S.P. & Buckley, D.L. (2012). Tracer kinetic modelling in MRI: Estimating perfusion and capillary permeability. *Physics in Medicine and Biology*, **57**, R1, doi:10.1088/0031-9155/57/2/R1.

Sourbron, S.P. & Buckley, D.L. (2013). Classic models for dynamic contrast-enhanced MRI. *NMR in Biomedicine*, **26**, 1004–1027, doi:10.1002/nbm.2940.

Steklova, K. & Haber, E. (2017). Joint hydrogeophysical inversion: state estimation for seawater intrusion models in 3D. *Computational Geosciences*, **21**, 75–94, doi:10.1007/s10596-016-9595-y.

Steuperaert, M., Debbaut, C., Carlier, C., Wever, O.D., Descamps, B., Vanhove, C., Ceelen, W. & Segers, P. (2019). A 3D CFD model of the interstitial fluid pressure and drug distribution in heterogeneous tumor nodules during intraperitoneal chemotherapy. *Drug Delivery*, **26**, 404–415, doi:10.1080/10717544.2019.1588423.

Stewart, G.W. (1993). On the early history of the singular value decomposition. *SAIM Review*, **32**, 551–566.

Stikov, N., Trzasko, J. & Bernstein, M. (2019). Reproducibility and the future of MRI research. *Magn Reson Med*, **82**, 1981–1983, doi:10.1002/mrm.27939.

Strang, G. (1980). *Linear algebra and its applications*. Academic Press, Inc., 2nd edn.

Studholme, C., Hill, D. & Hawkes, D. (1996). Automated 3-D registration of MR and CT images of the head. *Medical Image Analysis*, **1**, 163–175, doi:10.1016/S1361-8415(96)80011-9.

Taheri, S., Gasparovic, C., Huisa, B., Adair, J., Edmonds, E., Prestopnik, J., Grossetete, M., Shah, N., Wills, J., Qualls, C. & Rosenberg, G. (2011). Blood–Brain Barrier Permeability Abnormalities in Vascular Cognitive Impairment. *Stroke*, **42**, 2158–2163, doi:10.1161/STROKEAHA.110.611731.

Teo, Q., Thng, C., Koh, T. & Ng, Q. (2014). Dynamic Contrast-enhanced Magnetic Resonance Imaging: Applications in Oncology. *Clinical Oncology*, **26**, e9–e20, doi:https://doi.org/10.1016/j.clon.2014.05.014.

Thacker, N.A., Scott, M.L.J. & Jackson, A. (2003). Can dynamic susceptibility contrast magnetic resonance imaging perfusion data be analyzed using a model based on directional flow? *J. Magn. Reson. Imaging*, **17**, 241–255, doi:10.1002/jmri.10240.

Thomas, A., Arevalo-Perez, J., Kaley, T., Lyo, J., Peck, K., Shi, W., Zhang, Z. & Young, R. (2015). Dynamic contrast enhanced T1 MRI perfusion differentiates pseudoprogression from recurrent glioblastoma. *J Neurooncol*, **125**, 183–190, doi:10.1007/s11060-015-1893-z.

Thorup, L., Hjortdal, A., Boedtkjer, D.B., Thomsen, M.B. & Hjortdal, V. (2023). The transport function of the human lymphatic system—a systematic review. *Physiological Reports*, **11**, doi:10.14814/phy2.15697.

Tofts, P.S. (1997). Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. *Journal of Magnetic Resonance Imaging*, **7**, 91–101, doi:10.1002/jmri.1880070113.

Tofts, P.S. & Kermode, A.G. (1991). Measurement of the blood-brain barrier permeability and leakage space using dynamic mr imaging.

1. fundamental concepts. *Magnetic Resonance in Medicine*, **17**, 357–367, doi:10.1002/mrm.1910170208.

Tofts, P.S., Brix, G., Buckley, D.L., Evelhoch, J.L., Henderson, E., Knopp, M.V., Larsson, H.B., Lee, T.Y., Mayr, N.A., Parker, G.J., Port, R.E., Taylor, J. & Weisskoff, R.M. (1999). Estimating kinetic parameters from dynamic contrast-enhanced T1- weighted MRI of a diffusable tracer: Standardized quantities and symbols. *Journal of Magnetic Resonance Imaging*, **10**, 223–232, doi:10.1002/(SICI)1522-2586(199909)10:3<223::AID-JMRI2>3.0.CO;2-S.

van Dijken, B.R., van Laar, P.J., Smits, M., Dankbaar, J.W., Enting, R.H. & van der Hoorn, A. (2019). Perfusion MRI in treatment evaluation of glioblastomas: Clinical relevance of current and future techniques. *Journal of Magnetic Resonance Imaging*, **49**, 11–22, doi:10.1002/jmri.26306.

van Geuns, R.J.M., Wielopolski, P.A., de Bruin, H.G., Rensing, B.J., van Ooijen, P.M., Hulshoff, M., Oudkerk, M. & de Feyter, P.J. (1999). Basic principles of magnetic resonance imaging. *Progress in Cardiovascular Diseases*, **42**, 149–156, doi:10.1016/S0033-0620(99)70014-9.

van Herten, R.L., Chiribiri, A., Breeuwer, M., Veta, M. & Scannell, C.M. (2022). Physics-informed neural networks for myocardial perfusion MRI quantification. *Medical Image Analysis*, **78**, 102399, doi:10.1016/j.media.2022.102399.

van Houdt, P.J., Ragunathan, S., Berk, M., Ahmed, Z., Kershaw, L., Gurney-Champion, O., Tadimalla, S., Kallehauge, J., Arvidsson, J., Dickie, B., Lévy, S., Bell, L., Sourbron, S. & Thrippleton, M.J. (2022). Open Science Initiative for Perfusion Imaging (OSIPI): A community-led, open-source code library for analysis of DCE/DSC-MRI. *Proc. Intl. Soc. Mag. Reson. Med.*, 2691–2691.

van Houdt, P.J., Ragunathan, S., Berks, M., Ahmed, Z., Kershaw, L.E., Gurney-Champion, O.J., Tadimalla, S., Arvidsson, J., Sun,

Y., Kallehauge, J., Dickie, B., Lévy, S., Bell, L., Sourbron, S. & Thrippleton, M.J. (2023). Contrast-agent-based perfusion MRI code repository and testing framework: ISMRM Open Science Initiative for Perfusion Imaging (OSIPI). *Magnetic Resonance in Medicine*, doi:10.1002/mrm.29826.

van Osch, M.J., jan P.A. Vonken, E., Wu, O., Viergever, M.A., van der Grond, J. & Bakker, C.J. (2003). Model of the human vasculature for studying the influence of contrast injection speed on cerebral perfusion MRI. *Magnetic Resonance in Medicine*, **50**, 614–622, doi:10.1002/mrm.10567.

van Osch, M.J., Teeuwisse, W.M., Chen, Z., Suzuki, Y., Helle, M. & Schmid, S. (2018). Advances in arterial spin labelling MRI methods for measuring perfusion and collateral flow. *Journal of Cerebral Blood Flow Metabolism*, **38**, 1461–1480, doi:10.1177/0271678X17713434.

van Rijswijk, C.S.P., Hogendoorn, P.C.W., Taminiau, A.H.M. & Bloem, J.L. (2001). Synovial sarcoma: dynamic contrast-enhanced MR imaging features. *Skeletal Radiology*, **30**, 25–30, doi:10.1007/s002560000295.

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., İlhan Polat, Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., Vijaykumar, A., Bardelli, A.P., Rothberg, A., Hilboll, A., Kloeckner, A., Scopatz, A., Lee, A., Rokem, A., Woods, C.N., Fulton, C., Masson, C., Häggström, C., Fitzgerald, C., Nicholson, D.A., Hagen, D.R., Pasechnik, D.V., Olivetti, E., Martin, E., Wieser, E., Silva, F., Lenders, F., Wilhelm, F., Young, G., Price, G.A., Ingold, G.L., Allen, G.E., Lee, G.R., Audren, H., Probst, I., Dietrich, J.P., Silterra, J., Webber, J.T., Slavič, J., Nothman, J., Buchner, J., Kulick, J., Schönberger, J.L.,

de Miranda Cardoso, J.V., Reimer, J., Harrington, J., Rodríguez, J.L.C., Nunez-Iglesias, J., Kuczynski, J., Tritz, K., Thoma, M., Newville, M., Kümmerer, M., Bolingbroke, M., Tartre, M., Pak, M., Smith, N.J., Nowaczyk, N., Shebanov, N., Pavlyk, O., Brodtkorb, P.A., Lee, P., McGibbon, R.T., Feldbauer, R., Lewis, S., Tygier, S., Sievert, S., Vigna, S., Peterson, S., More, S., Pudlik, T., Oshima, T., Pingel, T.J., Robitaille, T.P., Spura, T., Jones, T.R., Cera, T., Leslie, T., Zito, T., Krauss, T., Upadhyay, U., Halchenko, Y.O. & Vázquez-Baeza, Y. (2020). Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature Methods*, **17**, 261–272, doi:10.1038/s41592-019-0686-2.

Wake, N., Chandarana, H., Rusinek, H., Fujimoto, K., Moy, L., Sodickson, D.K. & Kim, S.G. (2018). Accuracy and precision of quantitative DCE-MRI parameters: How should one estimate contrast concentration? *Magnetic Resonance Imaging*, **52**, 16–23, doi:10.1016/j.mri.2018.05.007.

Wang, S., Yu, X. & Perdikaris, P. (2020). When and why PINNs fail to train: A neural tangent kernel perspective. *Journal of Computational Physics*, **449**, doi:10.48550/arxiv.2007.14527.

Whitaker, S. (1986). Flow in porous media I: A theoretical derivation of darcy's law. *Transport in Porous Media*, **1**, 3–25, doi:10.1007/BF01036523.

Wicherts, J., Veldkamp, C., Augusteijn, H., Bakker, M., van Aert, R. & van Assen, M. (2016). Degrees of Freedom in Planning, Running, Analyzing, and Reporting Psychological Studies: A Checklist to Avoid p-Hacking. *Front Psychol*, **7**, doi:10.3389/fpsyg.2016.01832.

Willats, L. & Calamante, F. (2013). The 39 steps: evading error and deciphering the secrets for accurate dynamic susceptibility contrast MRI. *NMR in Biomedicine*, **26**, 913–931, doi:10.1002/nbm.2833.

Wilson, G., Aruliah, D.A., Brown, C.T., Hong, N.P.C., Davis, M., Guy, R.T., Haddock, S.H.D., Huff, K.D., Mitchell, I.M., Plumbley, M.D.,

Waugh, B., White, E.P. & Wilson, P. (2014). Best Practices for Scientific Computing. *PLoS Biology*, **12**, e1001745, doi:10.1371/journal.pbio.1001745.

Wilson, G., Bryan, J., Cranston, K., Kitzes, J., Nederbragt, L. & Teal, T.K. (2017). Good enough practices in scientific computing. *PLOS Computational Biology*, **13**, e1005510, doi:10.1371/journal.pcbi.1005510.

Woodall, R.T., Barnes, S.L., Ii, D.A.H., Sorace, A.G., Quarles, C.C. & Yankeelov, T.E. (2018). The Effects of Intravoxel Contrast Agent Diffusion on the Analysis of DCE-MRI Data in Realistic Tissue Domains. *Magn Reson Med*, **80**, 330–340, doi:10.1002/mrm.26995.

Wu, C., Hormuth, D.A., Oliver, T.A., Pineda, F., Lorenzo, G., Karczmar, G.S., Moser, R.D. & Yankeelov, T.E. (2020). Patient-Specific Characterization of Breast Cancer Hemodynamics Using Image-Guided Computational Fluid Dynamics. *IEEE Transactions on Medical Imaging*, **39**, 2760–2771, doi:10.1109/TMI.2020.2975375.

Wu, C., Hormuth, D.A., Easley, T., Eijkhout, V., Pineda, F., Karczmar, G.S. & Yankeelov, T.E. (2021). An in silico validation framework for quantitative DCE-MRI techniques based on a dynamic digital phantom. *Medical Image Analysis*, **73**, 102186, doi:10.1016/j.media.2021.102186.

Yushkevich, P., Gao, Y. & Gerig, G. (2016). ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images. *Annu Int Conf IEEE Eng Med Biol Soc*, 3342–3345, doi:10.1109/EMBC.2016.7591443.

Zachos, T.A., Aiken, S.W., DiResta, G.R. & Healey, J.H. (2001). Interstitial Fluid Pressure and Blood Flow in Canine Osteosarcoma and Other Tumors. *Clinical Orthopaedics and Related Research*, **385**, 230–236, doi:10.1097/00003086-200104000-00034.

Zhang, N., Zhang, L., Qiu, B., Meng, L., Wang, X. & Hou, B. (2012). Correlation of volume transfer coefficient Ktrans with histopathologic grades of gliomas. *J Magn Reson Imaging*, **36**, 355–363, doi:10.1002/jmri.23675.

Zhang, Q., Spincemaille, P., Drotman, M., Chen, C., Eskreis-Winkler, S., Huang, W., Zhou, L., Morgan, J., Nguyen, T.D., Prince, M.R. & Wang, Y. (2022). Quantitative transport mapping (QTM) for differentiating benign and malignant breast lesion: Comparison with traditional kinetics modeling and semi-quantitative enhancement curve characteristics. *Magnetic Resonance Imaging*, **86**, 86–93, doi:10.1016/j.mri.2021.10.039.

Zhang, Q., Luo, X., Zhou, L., Nguyen, T.D., Prince, M.R., Spincemaille, P. & Wang, Y. (2023). Fluid Mechanics Approach to Perfusion Quantification: Vasculature Computational Fluid Dynamics Simulation, Quantitative Transport Mapping (QTM) Analysis of Dynamics Contrast Enhanced MRI, and Application in Nonalcoholic Fatty Liver Disease Classification. *IEEE Transactions on Biomedical Engineering*, **70**, 980–990, doi:10.1109/TBME.2022.3207057.

Zhang, X., Petersen, E.T., Ghariq, E., Vis, J.B.D., Webb, A.G., Teeuwisse, W.M., Hendrikse, J. & van Osch, M.J.P. (2013). In vivo blood T_1 measurements at 1.5 T, 3 T, and 7 T. *Magn Reson Med*, **70**, 1082–1086, doi:10.1002/mrm.24550.

Zhou, L., Zhang, Q., Spincemaille, P., Nguyen, T.D., Morgan, J., Dai, W., Li, Y., Gupta, A., Prince, M.R. & Wang, Y. (2021). Quantitative transport mapping (QTM) of the kidney with an approximate microvascular network. *Magnetic Resonance in Medicine*, **85**, 2247–2262, doi:10.1002/mrm.28584.

Zhu, X., Li, K. & Jackson, A. (2005). *Dynamic Contrast-Enhanced MRI in Cerebral Tumours*, 117–143. Springer Berlin Heidelberg, doi:10.1007/3-540-26420-5_9.